

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:54:53 ; Search time 69.02 seconds  
(without alignments)  
793.551 Million cell updates/sec

Title: US-09-701-586B-2  
Perfect score: 2998  
Sequence: 1 MAARRRRSTGGGRARALNES.....PQVRRMYLLKQVFNFLQLW 570

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1176	39.2	635	2 T01311	NAD+ ADP-ribosyltr
2	1149.5	38.3	653	2 T03656	probable NAD+ ADP-
3	1022	34.1	1011	1 JH0581	NAD+ ADP-ribosyltr
4	1017	33.9	1013	1 S04200	NAD+ ADP-ribosyltr
5	1012.5	33.8	1014	1 A29725	NAD+ ADP-ribosyltr
6	1009	33.7	1016	1 S0428	NAD+ ADP-ribosyltr
7	1006.5	33.6	500	2 S26057	NAD+ ADP-ribosyltr
8	984.5	32.8	998	2 S31735	NAD+ ADP-ribosyltr
9	979.5	32.7	994	1 A47474	NAD+ ADP-ribosyltr
10	966	32.2	983	2 T51353	NAD+ ADP-ribosyltr
11	966	32.2	1009	2 C84719	probable poly (ADP
12	964	32.2	996	1 S42208	NAD+ ADP-ribosyltr
13	961.5	32.1	969	2 T03657	NAD+ ADP-ribosyltr
14	885.5	22.9	727	2 T18600	hypothetical prote
15	674.5	22.5	459	2 T08713	NAD+ ADP-ribosyltr
16	526	17.5	538	2 T20414	hypothetical prote
17	331.5	11.1	135	2 P80494	NAD+ ADP-ribosyltr
18	165.5	5.5	2004	2 D88948	protein ZK1005.1 [
19	141.5	4.7	181	2 T03058	NAD+ ADP-ribosyltr
20	134.5	4.5	1170	2 A72287	hypothetical prote
21	133	4.4	805	2 S48411	SEC6 protein - yea
22	130.5	4.4	1005	2 A64465	hypothetical prote
23	126	4.2	880	2 F75103	conserved hypotet
24	126	4.2	2058	2 A59267	myosin X - human
25	125	4.2	1051	2 T18351	Impl protein - Myc
26	125	4.2	1365	2 T30822	Impl protein - Myc
27	122.5	4.1	2401	2 T28676	rhoptery protein -
28	122	4.1	881	2 I84737	kinesin heavy chai
29	121	4.0	631	2 JC4298	hyaluronan recepto

30 120.5 4.0 1270 2 T22615 hypothetical prote  
31 120 4.0 1119 2 B70126 surface-located me  
32 120 4.0 3259 1 A56539 giantin - human  
33 119 4.0 1199 2 T29145 hypothetical prote  
34 119 4.0 1780 2 T12722 hypothetical prote  
35 118.5 4.0 1325 2 T42722 male-enhanced anti  
36 117.5 3.9 1940 2 A29320 myosin heavy chain  
37 117.5 3.9 2231 2 D71870 hypothetical prote  
38 117 3.9 963 1 A41919 kinesin heavy chai  
39 116.5 3.9 725 1 JC5016 hyaluronan recepto  
40 116.5 3.9 886 2 H69378 conserved hypotet  
41 116.5 3.9 1827 2 T16270 hypothetical prote  
42 116 3.9 504 2 S28298 hypothetical prote  
43 116 3.9 578 2 T40984 transcription fact  
44 115.5 3.9 540 2 S38085 transcripction fact  
45 115.5 3.9 600 2 A33658 paraflagellar rod

ALIGNMENTS

RESULT 1

T01311  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana  
N:Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T01311; S65662  
R:Kalicki, J.; Elliott, G.; Cloud, J.  
A:Description: The sequence of A. thaliana T14P8.  
A:Reference number: Z14290  
A:Accession: T01311  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-635 <KAL>  
A:Cross-references: EMBL:AF069298; NID:g3193282; PID:g3193299  
A:Experimental source: cultivar Columbia  
R:Lepiniec, L.; Babiychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.  
FEBS Lett. 364, 103-108, 1995  
A:Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(AD  
A:Reference number: S65662; MUID:95269779  
A:Accession: S65662  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-115; 'Gr', 116-635 <LEP>  
A:Cross-references: EMBL:Z48243; NID:g853721; PIDN:CAA88288.1; PID:g853722  
C:Genetics:  
A:Gene: PARP  
A:Map position: 4  
A:Introns: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3  
A:Note: T14P8.19  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

Query Match 39.2%; Score 1176; DB 2; Length 635;  
Best Local Similarity 43.9%; Pred. No. 9.9e-67;  
Matches 254; Conservative 98; Mismatches 187; Indels 40; Gaps 11;

QY 2 AARRRSTGGGR----ARALNESKRVNNGTAPEDSSPAKTRRCORQESKKMPVAGSKA 57  
DB 81 AIKGLDTTGTGKOLLERLCNDANNVSN---APVKSS-----NDEA 118  
QY 58 NKDRT---EDKQDESVALLLKGRAPDPECTAKV-GKAHYVCEGNDYVYMLNQTNIQF 113  
DB 119 EDDNNGFEERKEEKIVTATKGAVALDQWIPDEIKSYHVLRQDDVDYDAILNQTNRVD 178  
QY 114 NNNKYLIQLLEDQAQNFVSMWRGVRGKMGQHSLVACSNLNAKEIFOKKFLDKTKN 173  
DB 179 NNNFFVLQVLESCKKTYMYITRWGRVGRVGRGQSKLDGPDYSDWRRAIEFTNKFNDKTKN 238  
QY 174 NWEDREKFEKVPKGYDMLQMDYATNTQDEETKEESLSPKPE-SOLDLRVQELIKLI 232





N-Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib



Db 1008 KNEFKTSLW 1016

## RESULT 7

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 18-Jun-1999  
C:Accession: S26057; S78453; I52331  
R:Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.; Biochem. Cell Biol. 67, 653-660, 1989  
A:Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domain  
A:Reference number: I52331; MUID:90027702  
A:Accession: S26057  
A:Molecule type: mRNA  
A:Residues: 1-500 <THI>  
A:Cross-references: EMBL:X5497; NID:g56849; PIDN:CAA46478.1; PID:g56850  
R:Potvin, F.  
submitted to the EMBL Data Library, March 1992  
A:Reference number: S78453  
A:Accession: S78453  
A:Molecule type: mRNA  
A:Residues: 1-124, 'H', 126-127, 'A', 129-238, 'D', 240-500 <POT>  
A:Cross-references: EMBL:X5497  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match 33.6%; Score 1006.5; DB 2; Length 500;  
Best Local Similarity 42.6%; Pred. No. 3.7e-56;  
Matches 220; Conservative 101; Mismatches 168; Indels 27; Gaps 13;

QY 64 DKODESVKALLKGAQVDPDECTAKVGAHVYCGNDVYDMLNQTNLQFNNNKYYLIQL 123  
DB 3 NKSEKRMK-LTLKGAAVDPD-SGLEHSAHVLEKGGKVFSAITGLVDIVGTNSYYKLQL 60  
QY 124 LEDDAQNFVMMRWGRVGMKGQHSVLVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEK 183  
DB 61 LESDKESRYWFRSGRVGTVIGSKLEQMPKSEDAVEHFMKLYEERTGNNAHSHK-NFTK 119  
QY 184 VPKYDMLQDYATNTQDEETKEESLKSPLRP--ESQDLRLVQELIKLIGNVOAMEEM 241  
DB 120 YPKFYPLEIDYG---QDEAVK-----KLAVRPGTKSLPKPVQELVGMIFDVESMKKA 171  
QY 242 MMEKNTKAPLKLVAQIKAGYQSLKKTIEDCIRAGQHGRLMEACNEFYTRIPHDFG 301  
DB 172 LVEVEIDLQKMPGLKLSRQIQAYSILSEVQQAQVSGSSESQLLDLSNRYTLIPHDFG 231  
QY 302 LRTPLRTQKELSEKIQLEALGDIEIAIKLVK-TLQSPHPLDQHYRNHLCAALPLD 360  
DB 232 MKRPLLNNTDSVQAKVEMLDNLIDIEVAYSLLRGSSDDSKDPIDVNYEKLKTDIKVVD 291  
QY 361 HESYEFKVISQYLOSTHAPTHSDYTWTLTLLDFEVEKDGEKEAFR--EDLNRLMLLHWSR 418  
DB 292 RDSSEAEVIRKYVKNTHATTHNAYDLEVIDIFKIERGESQRYKPFQNLNRLHWSR 351  
QY 419 MSNMGVLSHGLRIAPPEAPITGYMFGKGYFADMSKSNANYCFASRLKNTGLLLSEVA 478  
DB 352 TTNFAGILSOGLRIAPPEAPVGYMFGKGYFADMSKSNANYCHTSQGDPIGLILGEVA 411  
QY 479 LGQCNELLENAPKAEGLQKHSKTLGKMAPSSAHFVTLNGSTVPLGASDGTILNPDG 538  
DB 412 LGNMYELKHASHISK-LPKGKHSVGLGKTAPDPSASITLDGEVPLG---TGI--PSG 464  
QY 539 YT---LNYNEYIYVNPQVRMYLLKVOFNF-LQLW 570  
DB 465 VNDTCLLIYNEYIYVNDIAQVNLKYLKLLKFNKFTSLW 500

## RESULT 8

S31735  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)

N:Alternate names: poly ADP-ribose polymerase  
C:Species: Xenopus laevis (African Clawed frog)  
C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Jun-2000  
C:Accession: S31735; PN0495  
R:Saulier-le Drian, B.M.  
submitted to the EMBL Data Library, May 1992  
A:Reference number: S31735  
A:Accession: S31735  
A:Molecule type: mRNA  
A:Residues: 1-998 <SAU>  
A:Cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PID:g1334661  
R:Ozawa, Y.; Uchida, K.; Uchida, M.; Ami, Y.; Kishida, S.; Okada, N.; Miwa, M. Biochem. Biophys. Res. Commun. 193, 119-125, 1993  
A:Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose) polymerase  
A:Reference number: PN0494; MUID:93277538  
A:Accession: PN0495  
A:Molecule type: mRNA  
A:Residues: 742-745, 'E', 747-876 <OZA>  
C:Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and diff  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pento

Query Match 32.8%; Score 984.5; DB 2; Length 998;  
Best Local Similarity 39.1%; Pred. No. 2.4e-54;  
Matches 222; Conservative 109; Mismatches 196; Indels 41; Gaps 14;

QY 21 KRYVNGNTAPE-----DSSPAKTRCQROESKKMPVAG-----GKANKORTEDKQD 67  
DB 444 KEVSGKSVQELLSQFGISSWGAIEKQEAQVOTPEKSSGPGVAGKSSGKKEEKNKSE 503  
QY 68 ESKVALLKKGKAPVDPCTAKVGAHVYCGNDVYDMLNQTNLQFNNNKYYLIQLLEDD 127  
DB 504 KKK-LTVKGGAIDPDESELE-DSCHVLETTGGKIFSATLGLVDITRGTNSYYKLQLEHD 561  
QY 128 AQNFVMMRWGRVGMKGQHSVLVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVP 186  
DB 562 RDSRYWFRSGRVGTVIGSKLEEMSSK-EDATEHPLNLYQDKTGNNAHSH-PNFTKYPK 619  
QY 187 KYDMLQDYATNTQDEETKEESLKSPLRPESQDLRLVQELIKLIGNVOAMEEMEMK 246  
DB 620 KFYPLEIDYG---QEDDVKK---LSVAGTKSLKAPVQELIKLIPDVESMKKAMVEFE 673  
QY 247 YNTKKAPLGLTVAQIKAGYQSLKKTIEDCIRAGQHGRLMEACNEFYTRIPHDFGLRTP 306  
DB 674 IDLQKMPGLKLSRQIQAYSILSQQAQVSESLSSEARLLDLSNQFYTLIPHDFGKMKPP 733  
QY 307 LIRTQKELSEKIQLEALGDIEIAIKLVKTELQSPH-HPLDQHYRNHLCAALPLDHESEY 365  
DB 734 LLANLEYIQAKVQMLDNLIDIEVAYSLLRGSGADDGKDPIDVYKYEIKTDIKVYAKDSEE 793  
QY 366 FKVISQYLOSTHAPTHSDYTWTLTLLDFEVEKDGEKEAFR--EDLNRLMLLHWSRMSNV 423  
DB 794 SRICDVTKNTHATTHNAYDLEVIDIFKIERGEYKPKFQNLNRLHWSRRTNFA 853  
QY 424 GILSHGLRIAPPEAPITGYMFGKGYFADMSKSNANYCFASRLKNTGLLLSEVALQCN 483  
DB 854 GILSGLRIAPPEAPVGYMFGKGYFADMSKSNANYCHAMPSPGILGILLGEVALGNMH 913  
QY 484 ELLENAPKAEGLQKHSKTLGKMAPSSAHFVTLNGSTVPLGPA-----SDTGILNPDG 538  
DB 914 E-LKAASQITKLPKHSVGLGKTAPDPSATVQLQGDVYPLGKGTSAINTSDTSL---- 968  
QY 539 YTLNNEYIYVNPQVRMYLLKVOFNF 566  
DB 969 ----YNEYIYVNDIAQVNLKYLKLLKFN 992

## RESULT 9

A47474

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000



A:Molecule type: DNA  
A:Residues: 1-1009 <STO>  
A:Cross-references: GB:AE002093; NID:g4432827; PIDN:AA020677.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g31320  
A:Map position: 2  
C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 32.2%; Score 966; DB 2; Length 1009;  
Best Local Similarity 38.8%; Pred. No. 3.5e-53;  
Matches 213; Conservative 115; Mismatches 151; Indels 70; Gaps 17;

QY 45 QESKMPVAGGKANKDRTEKQDESVAKALLLKAPVDPDECTAKYGVCEGNDVYDV 104  
DB 498 KKORKLPF-----DKYIED-TSESLVTVKVGSAVH-EASGLQEHCHILEDGNSIYNT 550  
QY 105 MLNQTNLQFNKKYLIQLLEDAAQDNFSVWRWGRVG--KMGQHSVLVACSNLNAKEI 162  
DB 551 TUSMSDLSTGINSYVLIQIEDKSGDCYVFRKWRGVNEKIGG-----NKVEEM 600  
QY 163 -----FQKFLDKTKNNWEDREK---FEKVPCKYDMLQMDYATNTQDEETKKEESL 211  
DB 601 SKSDAVHEKRLFLKKTGNTWESWQKTFNQKQKFLPLDIDYGN-----KQVAK 652  
QY 212 KSLPKPESOLDLRVOELIKLICNVQAMEEMMEMKYNKTKAPLGKLTVAQIKAGYQSLKK 271  
DB 653 KEPTQSSNAPLSLTELKMLFDVETYSAMMEFEINNSEMPGLKSKHNIQKGFALTE 712  
QY 272 IEDCI-----RAGQGRALMEACNEFFYTRIP--HDFGLRTPLIQTOKELSEKIQLLEAL 324  
DB 713 IQRLETSDPQPTMKESLVDASNFFTPISIH-----PHIIRDEDDFKSKVKMLEAL 766  
QY 325 GDIETAIKVLTKLOSPHPLDOHVRNLHLCALPLDHDSEYEFKYSQYLQSTHAPTHSDY 384  
DB 767 QDIEASRIVGPDVSTE-SLDDKYTKLHCDLSPHOSDYLKTEKLYNLTHTHAPTHSEW 825  
QY 385 TMTLDDLFEVEKDGKEAF---REDLHNRMLLHGRSMNVWVGLSHGLRIAPPEAPITG 441  
DB 826 SLELEEVFALEREGEFDKYAPHREKLGKMLLHWSRLTNFVGLNQGLRIAPPEAPATG 885  
QY 442 YNFGGIYFADMSKSNKYCFCASRLKNTGTLNLSVALGQCNELLEAN-----PKAEGLL 496  
DB 886 YNFGGIYFADLVKSAQYCYCTCKKNPVGLMLLSVALGEIHELTKAKYMDKPPR----- 940  
QY 497 QGKHSYTKGLGKMAPSSAHFVTLNGS--TVPLGPASDTGILNPDGYTLNXYVNPNOVR 555  
DB 941 -GKHSYTKGLGKVPQDSEFAKWRGDTVPCKGKPVSKVASE---LMTNXYIYDITQVK 996  
QY 556 MRYLLKVQF 564  
DB 997 LQFLKVRP 1005

RESULT 12  
S42208  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)  
A:Alternate names: poly(ADP-ribose) polymerase  
C:Species: Sarcophaga peregrina  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: S42208; S71496  
R:Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat  
Eur. J. Biochem. 220, 607-614, 1994  
A:Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarc  
A:Reference number: S42208; MUID:94110813  
A:Accession: S42208  
A:Molecule type: mRNA  
A:Residues: 1-996 <MAS>  
A:Cross-references: EMBL:D16482; NID:g473742; PIDN:BA03943.1; PID:g538248  
A:Accession: S71496  
A:Molecule type: protein  
C:Superfamily: NAD+ ADP-ribosyltransferase

C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger  
F:1-369/Domain: DNA binding #status predicted <DNA>  
F:370-507/Domain: auto-modification #status predicted <AMO>  
F:508-996/Domain: NAD binding #status predicted <NAD>

Query Match 32.2%; Score 964; DB 1; Length 996;  
Best Local Similarity 39.9%; Pred. No. 4.7e-53;  
Matches 216; Conservative 116; Mismatches 177; Indels 32; Gaps 13;

QY 33 SSPAKTRCQQRQESKKMPVAGGKANKDRTEKQDESVAKALLLKAPVDPDECTAKYVKA 92  
DB 480 TDPATRITQESKSKSIYTKSPKSMT-----LKIKDLAVDPD-SGLEDAV 528  
QY 93 HVCBGN-DVYDVMNLQTNLQFNKKYLIQLLEDAAQDNFSVWRWGRVG-KMGQHSILV 150  
DB 529 HYYVSRNEKYNVLGIDIDIKNNKNSFYKLQLESMDKNNRFFVFRSWRGITGTTGNKLD 588  
QY 151 ACSGNLNAKEIFQKFLDKTKNNWEDREKFEKVPCKYDMLQMDYATNTQDEETKKEES 210  
DB 589 NFS-NLVDATIVQFKELYLEKSGNHFPENRENFVKVAGRWYPIDIDYA-----EDSKIDLS 641  
QY 211 LKSPKLPESOLDLRVOELIKLICNVQAMEEMMEMKYNKTKAPLGKLTVAQIKAGYQSLK 270  
DB 642 AEHDIK--SKPLSVQDIIKLMFDVDSMKRTMMEFDLDMKMLPLGKLSOKQIQSAKYVLT 699  
QY 271 KIEDCIRAGQGRALMEACNEFFYTRIPHDGFLRTPLIQTOKELSEKIQLLEALGDIETA 330  
DB 700 EYELIQGGTNAKFI DATNRFYTLIPHNFQTSQSPPLDTPTEQVQLRQMLDLSLEICA 759  
QY 331 IKLVKTELOSPHPLDOHVRNLHLCALPLDHDSEYEFKYSQYLQSTHAPTHSDYTMTL 389  
DB 760 YSLQTESKADINFDKHYEQLTKEPLDKNSEYILLQKYVKNTHAETHKLYDLEVV 819  
QY 330 DLFEVEKDGKEAF--EDLHNRMLLHGRSMNVWVGLSHGLRIAPPEAPITGTMFGKG 447  
DB 820 DIFKVARQGEARRYPFKKLNHRLHWSRLTNFAGILSHGLKIAPPEAPVGYTMFGKG 879  
QY 448 IYFADMSKSNKYCFCASRLKNTGTLNLSVALGQCNELLEANPKAEGLLQGHKSTPKLGK 507  
DB 880 IYFADMSKSNKYCCTSHHNSHTGLMLLSVALGDMMECTAAKYVTK-LPNKHSFCFGRGR 938  
QY 508 MAPS-SAHFVTLNGSTVPLG-PASDTGILNPDGYTLNXYVNPNOVRMYRLYLKVQFN 565  
DB 939 TMPNPSESIREDGVEIPLGKPIITNDSLKS-----SLLYNEFIYDIAQVNIQYMLRMNEK 994  
QY 566 F 566  
DB 995 Y 995

RESULT 13  
T03657  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) 2 - maize  
C:Species: Zea mays (maize)  
C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 20-Jun-2000  
C:Accession: T03657  
R:Babychuk, E.; Cottrill, P.; Storozhenko, S.; Fuangthong, M.; O'Farrell, M.; Van Mo  
submitted to the EMBL Data Library, November 1997  
A:Description: Higher plants possess two poly(ADP-ribose) polymerases.  
A:Reference number: Z14992  
A:Accession: T03657  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-969 <BAB>  
A:Cross-references: EMBL:AJ222589; PIDN:CAA10889.1  
C:Genetics:  
A:Gene: PARP2  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase  
Query Match 32.1%; Score 961.5; DB 2; Length 969;

```

79 APVDPECTAK-----VGKAHVYCEGNDVYDMLNQTNLQFNNNKYYLIQLLED 126
   || | | | : : : : : | : | : | : | : | : | : | : | : | : | :
243 GNPDENDFAKKRMKKEARLMVEQKRMMKQSOLLWEIQRIFERMPYTDN----ISLRE 298
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
127 DAORNFSVNMWRGVGMQHSLVACSGNLNKAKEIFOKFKFLDKTKNNWEDREKFVPKG 186
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
299 NEOD-----IPEGHDRTAQDVPH---EKTNDWIYRKHFRKMFG 334
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
187 KYDMLQMDYATNTQDEETKKEBSLKPLPESQ--LDLRVQELIKLICNVQAEEEMME 244
   : : || : : : : : : : : : : : : : : : : : : : : : : : : : :
335 MFSYVETDYSEFVTGNHKKKITPGSKITPGSKTLLPKSVKEVMISFDVENMKSAKS 394
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 MKYNTKAPGLKVIAQIRKAGYSCLKIED-CIRAGOHGRALMEACNEFFTRIPHDGLR 303
   : : ||| : | : | : | : | : : : : : : : : : : : : : : : : :
395 FEDVNKMPGLRSHQNINAFVLDNISDLLVKLPIDASKILLDFSNKFTIYPHNFMGR 454
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
304 TPLLRTQKELSEKIOLLEALGDIEIAKLIV-----KTQLSPEHPDLQHYNLHCAL 356
   | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
455 VPEIDSFHKIKEKNMNLALLDIKAYDOI SGDPVASTSLSI--DPVDINYRKLA CIM 512
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
357 RPLDHSEYEKVTISOYLSQTHAPTHSDYTMTLLDLFEVXDGEKEAPREDLNRMLLWHG 416
   || : : |||| : | : | : | : | : | : | : | : | : | : | : | :
513 EPLQGCGDDWNMIHQYLKNTFHGATH-DLKVELIDILKVNDRDNESKPKRIGNRRLLWHG 571
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 SRMSNVGILSHGLRTAPPEAITGYMGAGIYFADMSSKSNVCASFASRLKNTGLLLSE 476
   : : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
572 SGKMFAGILQQGLRIAPPEAVSGYMGVGFADMFSGSYFCVRAN-AKEEAYLLLCD 630
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
477 VALGOCNELLEA-NPKAEGLLQKHSHTGLGKMAPSS-AHFVTLNGSTVPLGPA--SDTG 532
   |||| : : | : | : | : | : | : | : | : | : | : | : | : | :
631 VALGNVQQLMASKNVSRQTLPAGFSQVOGVGROCPREIGSYKNPDPGTVPGLTYMQLOG 690
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
533 ILNPDCYTLNYEYIYNPNQVBRVYLLKYOFN 565
   | | | | | | | | | | : : : : : : : : : : : : : : : : : :
691 KQNVQ-YHLLYNEFIYVDVQIQLYLVVRKMH 722
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 15
T08713
NAD+ ADP-ribosyltransferase homolog DKFzp566G0224.1 - human (fragment)
C:Species: Homo sapiens (man)
C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 03-Nov-2000
C:Accession: T08713
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, May 1999
A:Reference number: Z16472
A:Accession: T08713
A:Molecule type: mRNA
A:Residues: 1-459 <ANS>
A:Cross-references: EMBL:AL050034
A:Experimental source: fetal kidney; clone DKFzp566G0224
C:Genetics:
A>Note: DKFzp566G0224.1
C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 22.5%; Score 674.5; DB 2; Length 459;
Best Local Similarity 35.1%; Pred. No. 3.3e-35;
Matches 170; Conservative 91; Mismatches 163; Indels 61; Gaps 15;

QY 111 LQFNKKYYLIQLLEDAQRNSFWMRGWGKMGHOSLVACSGNLNKAKEIFOKFKELD 170
   : : |||| : |||| : | : | : |||| : ||| : | : | : ||| : ||| : |
DB 1 IENNKKFYIIQLQ--DSNRFFTCWNRWGRVGGQ-SKINHTRLEDAKDFEKREK 58
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 171 TKNNWEDREKFEKVPQKYMLOMDYATNTODEBETKKEBSLKSPKPESQ-----LDLR 224
   |||| : | |||| : ||| : | : | : ||| : ||| : | : | : | : | :
DB 59 TKNNWAERDHFVSHPKYLTLEV---QAEDAQAQAVVYDVRGPVRTVTKRVQPCSLDPA 114
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 225 VQELIKLICNVQAEEEMENKNTTKAPGLKTVIAQIRKAGYSCLKIEDCIRA-GOHR 283
   : : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
DB 115 TQKLITNIFSKEFMKNTMALMDLDVKNMPLJGKLSKOQIARGFALEALRGALGPTDGGQ 174
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 284 ALMEACNEFFTRIPHDGLRTPPLIRTOKELSEKIOLEALGDIEI---ATKLKVLTELSP 341

```

```

:| | : ||| |||:| | | : | | | | | | : | : :
Db 175 SLEELSSHFTYVIPHNFHSGOPPPINSPELLQAKKMDLLVLADIELAQAVSEQEKTV 234
QY 342 E---HPLDOHYRNHLHCAALPLDHESYEFKVISQYLOSTHAPTHSDYTMILLDLFEYKDG 398
| ||||:| : | | : | | : | | : | | : | | : | | : | |
Db 235 EEPHPLDRDYQLKCOLQLDSDGAPEYKVOTYLEOTGS---NHRCPQLQHIWKYNQEG 291
QY 399 EKEAFR--EDLHNRMLLWHGSRMSNVVGIILSHGLRIAPPEAPITGYMFGKGIYFADMSSK 456
| : | : | | | | | : | : | | | | : | : | | | | : | |
Db 292 EEDRFOAHSKLGNRKLLWHGCTNMAVVAAILTSLRIMPH----SGGRVGEIGIYFASENSK 347
QY 457 SANCFASR--LKN TGILLLLSEVALQOCNELLEANPKAECGLQGHSTKGLGKMAPSSAH 514
| | : | : | | | | : | | : | | : | | : | | : | |
Db 348 SAGYVIGMCKGAHHVGMFLGEVALGREHHINTDNPSL-----KSPPPGFD 393
QY 515 FVTLNGSTVPLGPASDT-----GILNPDGY-----TLNYNEXIVYNPNQVRMR 557
| | | | | | | : | | : | | : | | : | | : | | : | |
Db 394 SVIARGHTEP-DPTQDTELELDGQQVVVFGQGPVPCPEFSSSTFSOSEYLLIYOESQCLRL 452
QY 558 YLLKV 562
| | | : |
Db 453 YLLEV 457
```

Search completed: August 29, 2002, 07:59:24  
Job time: 271 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: August 29, 2002, 07:59:29 ; Search time 69.02 Seconds

(Without alignments)  
751.786 Million cell updates/sec

Title: US-09-701-586b-6

Perfect score: 2854

Sequence: 1 MSLEFLAAMPKRPWVQTEG.....EYLIVQESQCRRLYLEVHL 540

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR-71:\*  
2: PIR1:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2419	84.8	459	2	T08713	NAD+ ADP-ribosyltr
2	694	24.3	635	2	T01311	NAD+ ADP-ribosyltr
3	681	23.9	1016	1	JS0428	NAD+ ADP-ribosyltr
4	674.5	23.6	996	1	S42208	NAD+ ADP-ribosyltr
5	670.5	23.5	653	2	T03656	Probable NAD+ ADP-
6	670	23.5	1014	1	A29725	NAD+ ADP-ribosyltr
7	669.5	23.5	1011	1	JH0581	NAD+ ADP-ribosyltr
8	657	23.0	500	2	S26057	NAD+ ADP-ribosyltr
9	653	22.9	998	2	S31735	NAD+ ADP-ribosyltr
10	653	22.9	1013	1	S04200	NAD+ ADP-ribosyltr
11	638	22.4	994	1	A47474	NAD+ ADP-ribosyltr
12	598.5	21.0	983	2	T51353	NAD+ ADP-ribosyltr
13	598.5	21.0	1009	2	C84719	Probable poly (ADP
14	578	20.3	969	2	T03657	NAD+ ADP-ribosyltr
15	534	18.7	727	2	T18600	hypothetical prote
16	393.5	13.8	538	2	T20414	NAD+ ADP-ribosyltr
17	190.5	5.8	135	2	PN0494	NAD+ ADP-ribosyltr
18	164.5	5.8	2004	2	D88948	protein ZK1005.1
19	134	4.7	1156	2	B70356	chromosome assembl
20	126	4.4	1222	2	C88504	protein B0361.3
21	120.5	4.2	2044	2	AB1180	probable peptidogl
22	120	4.2	1938	1	A40997	myosin heavy chain
23	116.5	4.1	1435	2	S69632	regulatory protein
24	113.5	4.0	181	2	T03058	NAD+ ADP-ribosyltr
25	113.5	4.0	1134	2	A60234	IgA Fc receptor pr
26	113.5	4.0	1164	1	FCSONG	IgA Fc receptor pr
27	113.5	4.0	4540	2	T30838	cytoplasmic dynein
28	113	4.0	2078	2	T25400	hypothetical prote
29	110	3.9	1534	2	A56734	ribosome receptor,

30	110	3.9	2288	2	T29999	hypothetical prote
31	109.5	3.8	446	1	A46335	gag polyprotein -
32	108.5	3.8	368	1	QOYV	transforming prote
33	108.5	3.8	1233	2	T30534	chromosome segrega
34	108.5	3.8	1871	2	S27938	hypothetical prote
35	108	3.8	359	2	T12540	hypothetical prote
36	108	3.8	1092	2	T33717	carbamoyl-phosphat
37	108	3.8	3225	2	I52300	giantin - human
38	108	3.8	3259	1	A56539	giantin - human
39	107.5	3.8	716	2	T26998	hypothetical prote
40	107	3.7	1166	2	G69708	chromosome segrega
41	106.5	3.7	568	2	T34522	hypothetical prote
42	106	3.7	725	1	JC5016	hyaluronan recepto
43	106	3.7	932	2	S62555	protoplast regener
44	106	3.7	1199	2	T29145	hypothetical prote
45	106	3.7	2176	2	T13806	toucan gene protei

## ALIGNMENTS

RESULT 1  
T08713  
NAD+ ADP-ribosyltransferase homolog DKFZp566G0224.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Nov-2000  
C:Accession: T08713  
R:Ansoorge, W.; Wilkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, May 1999  
A:Reference number: 216472  
A:Accession: T08713  
A:Molecule type: mRNA  
A:Residues: 1-459 <ANS>  
A:Cross-references: EMBL:AL050034  
A:Experimental source: fetal kidney; clone DKFZp566G0224  
C:Genetics:  
A>Note: DKFZp566G0224.1  
C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match	84.8%	Score 2419;	DB 2;	Length 459;
Best Local Similarity	99.8%	Pred. No. 1,4e-161;		
Matches 458;	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;
QY	82	TENNNNKFFYITQLLODSNREFTCNNRMRGVEGOSKINHTRLDPAKKDEKKFREXTK	141	
DB	1	TENNNNKFFYITQLLODSNREFTCNNRMRGVEGOSKINHTRLDPAKKDEKKFREXTK	60	
QY	142	NNMAERDHFVSHPGKYLIEVQAEDEAEVAVKVDGCFVTRRVOPCSLDPAATOKLIT	201	
DB	61	NNMAERDHFVSHPGKYLIEVQAEDEAEVAVKVDGCFVTRRVOPCSLDPAATOKLIT	120	
QY	202	NIFEKEMFKNTMALMDLVKKMPLGKLSKQOIAFGFALALBALKGPTDGGSGLEELS	261	
DB	121	NIFEKEMFKNTMALMDLVKKMPLGKLSKQOIAFGFALALBALKGPTDGGSGLEELS	180	
QY	262	SHEFTVTPHNFGSHQPPINSPPELLQAQKMLVLADIELAQAQAVSEOKTYEEVPHR	321	
DB	181	SHEFTVTPHNFGSHQPPINSPPELLQAQKMLVLADIELAQAQAVSEOKTYEEVPHR	240	
QY	322	LDROYQLKCOLQLDLSAPERYKIVQYLFQTSNHRCPYQHTMKVQEEDEDRFOHS	381	
DB	241	LDROYQLKCOLQLDLSAPERYKIVQYLFQTSNHRCPYQHTMKVQEEDEDRFOHS	300	
QY	382	KLGNRKLLMHGNTNNAVVAAILTSGLRIMPHSGRGVKGIGYPASENSAGYVIMKCGAH	441	
DB	301	KLGNRKLLMHGNTNNAVVAAILTSGLRIMPHSGRGVKGIGYPASENSAGYVIMKCGAH	360	
QY	442	HVGWTFGEVALGREHNTDNPISLSPPGEDSVIARNGHEPPTOTTELDGGOOVV	501	
DB	361	HVGWTFGEVALGREHNTDNPISLSPPGEDSVIARNGHEPPTOTTELDGGOOVV	420	
QY	502	PGQVPVPCPEFFSSSTFSQSEYLIQESQCRRLYLEVHL	540	

Db 421 PGQVPVPCPEFSSSTFSQSEXYLIQESQCRRLRYLLEVLH 459

RESULT 2

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana  
N:Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T01311; S65662  
R:Kallack, J.; Elliott, G.; Cloud, J.  
Submitted to the EMBL Data Library, May 1998  
A:Description: The sequence of A. thaliana T14P8.  
A:Reference number: 214290  
A:Accession: T01311  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
A:Residues: 1-635 <RAL>  
A:Cross-references: EMBL:AF069298; NID:93193282; PID:93193299  
A:Experimental source: cultivar Columbia  
R:Leplanc, L.; Babychuk, E.; Kushnir, S.; Van Montagu, M.; Inze, D.  
FEBS Lett. 364, 103-108, 1995  
A:Title: Characterization of an Arabidopsis thaliana cDNA homologue to animal poly(ADP-ribose)  
A:Reference number: S65662; M01D:95269779  
A:Accession: S65662  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-115, 'GT', 116-635 <LEP>  
A:Cross-references: EMBL:Z48243; NID:9853721; PID:9853722  
C:Genetics:  
A:Gene: PARP  
A:Map position: 4  
A:Insertions: 28/3; 116/1; 162/3; 190/1; 268/3; 312/1; 330/3; 355/2; 371/1; 387/3; 408/3; 4  
A:Note: T14P8.19  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger

Query Match 24.3%; Score 694; DB 2; Length 635;  
Best Local Similarity 34.2%; Pred. No. 1.2e-40;  
Matches 185; Conservative 97; Mismatches 221; Indels 38; Gaps 18;  
Db 107 SNAPVSSNDEADEDDNGFEKEEKETATKGAVALDQWIPDEIKSQYHVLQRGDDV 166  
QY 18 TEGPEKKKGROGREDEPFRSTAEALKAIPAERITRV-DPTCPSSNPTQYED--- 72  
Db 167 YDAILNQTNRDNNKFEVLQVLESQSKTYMYTTRWGRGVAGQSKLDGPYDSMDRAIE 226  
QY 73 YNCTINQNTENNKKRYITIOQLQ-DSNREFTCNRWGRGVEGQSKIN-HFRLEDAKK 130  
Db 167 YDAILNQTNRDNNKFEVLQVLESQSKTYMYTTRWGRGVAGQSKLDGPYDSMDRAIE 226  
QY 131 DFEKKFRETKNMMAERDHFVSHPGKTYTLIEVOADEAOEAVVVDGPRVTYTKRQP- 189  
Db 227 IFTNNKFNKTKNWSDKKEITPRKSTYTWLEMDYGREENDSPVND---IPSSSEYKPE 283  
QY 190 -GSLDPATOKLTNITFSKEMFKNTMALMDLVKMPKIGLSKQOIAFGPALEALEBALK 248  
Db 284 QSKLDRVAFILICNVSMACHMEIGYNAKPLPGKISKSTISYSGYEVLRKRISEVI- 342  
QY 249 GPRDGGQSEELSSHFTYVPHNFGHSQPP--INSPPELLAKMDLVADIELAQALQ 306  
Db 343 -DRYDRTLEELSGEFTYVPHDFGFKMSQFVITPQKIKOKIEVEALGEIETLAKKL 401  
QY 307 AVSEOEKTEVEVPHDRDYLQKCOLQDLDSGAREKVKVLOTLEOT-GSNHRCPTIQ-- 363  
Db 402 SVDPGQLQ----DDPLYHNGQLNCGITPYGNDSEESMANTMENHATHSGYVEIA 456  
QY 364 HIKVNOEGEEDRFQASHKIGNRLMHTGTMVAAILTSGLRIMPH---SGGRVKG 419  
Db 457 QLEFRASRAVEADRFQFSSSSKNNMLMHGSGRLTNMAGIILSQGLRIAPREAPVGYMGKG 516  
QY 420 IYVASNSKSGAGVYIGMKCAHHVGMFGEVNLGREGHINTNTNPISLAKSPRPEDSYIAR 479  
Db 517 VYFADMFSSANSYCYA-NTGAND-GVALLICEVALGDMNELLSDYMANDLNLPKRLSTKGV 574

QY 480 GHTPEPTQDTELELDGQVYVPGQPV--PCPEFSSSTFSQSEXYLIQESQCRRLRYLLE 537  
Db 575 GKTAPEPNSQAQTLIE-DG-VVVPFGKRVKSC---SKGMLLYNEYIYVNEQJMKRVYIQ 628

QY 538 Y 538

Db 629 V 629

RESULT 3

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine  
N:Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate)  
C:Species: Bos primigenius taurus (cattle)  
C>Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
C:Accession: J50428; S00328; A30458  
R:Salto, I.  
Submitted to JIPID, February 1990  
A:Reference number: J50428  
A:Accession: J50428  
A:Molecule type: mRNA  
A:Residues: 1-1016 <SAT>  
A:Experimental source: thymus  
R:Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T  
Eur. J. Biochem. 171, 571-575, 1988  
A:Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int  
A:Reference number: S00328; M01D:88151954  
A:Accession: S00328  
A:Molecule type: mRNA  
A:Residues: 648-714; 838-904 <TAN>  
A:Cross-references: EMBL:X06986  
A:Accession: A30458

A:Molecule type: protein  
A:Residues: 658-685; 689-696; 893-901 <TA2>  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loop  
F:21-51/Region: zinc finger  
F:128-165/Region: zinc finger  
F:200-220/Region: helix-turn-helix motif  
F:224-231/Region: nuclear location signal  
F:250-270/Region: helix-turn-helix motif  
F:294-301/Region: nucleotide-binding motif A (P-loop)  
F:890-903/Region: nucleotide binding #status predicted

Query Match 23.9%; Score 681; DB 1; Length 1016;  
Best Local Similarity 33.9%; Pred. No. 1.8e-39;  
Matches 188; Conservative 99; Mismatches 205; Indels 62; Gaps 24;

QY 7 AMAPKPK----PWVOTEGPEKKKGROGREDEPFRSTAEALKAIPAERITRVPTCPPLS 62  
Db 494 AVGPBKSGAAPSCKSKSGPYKEG--TNKSEKRMKLIKGA-----VDPQSGLE 542  
QY 63 SNPTQYED---YNCTINQNTENNKKRYITIOQLQDSNRF-FTCWNRWGVEV-GQ 116  
Db 543 HN--AHVLEKGKVFSAITGLVIDVKTNSYKILQLEDDEKESRYMIFRSWGHVYIGS 600  
QY 117 SKINHFRLDAKDKPEKKREKTKNNMAERDHFVSHPGKTYTLIEVQ-ADDEAOEAVVY 175  
Db 601 NKLQMSKEDALHMKLYEETGNAMWSK-NFTNHPKFFYLEIDYGDG--EAVKLI 657  
QY 176 DRPVRVTTRVPOCSLDPATOKLTNITFSKEMFKNTMALMDLVKMPKIGLSKQOIA 235  
Db 658 TVNP-GTKSLRPP-----VQNLIKMIFPEVSKKAMAYEETIDQKMPKIGLSKQOIA 710  
QY 236 GFEALEALEAL-KGPTDGGQSEELSSHFTYVPHNFGHSQPPINSPPELLQAKMDLL 294  
Db 711 AYSILSRVQALQSGSSD--SHILDSNRYTTLIPDFGKMPKPLNANNSVQAKYEMLD 768  
QY 295 VLADIELAQALQAVSEOEKTEVEVPHDRDYLQKCOLQDLDSGAREKVKVLOTLEOT- 353  
Db 769 NLDIEVAVSLRGGSDSSKD---PIDVNYEKLTKTDIKVVDKDESEAEIIRKYVKNTH 824



```

Oy 354 GSNHRCPTLQ--HTRKVNQGEDEDRFOASHKLGKRLMHGTMAVVAALITSGRLTMRP 411
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 825 ATTHNAYDLVLEVDLFKIEREESOSRYKRPKQOLMNRRLMHGSRTTNGFALISQGLTMRP 884
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 412 ----SGRWGKGIYFASNSKSGAGYVIGMKCGAHH---VGYMFLGEVVALGREHHINTDNP 464
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 885 EAPRYVGYMGFKGIYFADMYVSKSAN-----CHTSGDPRIGLILGLCEALGMMYTELKHAR- 938
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 465 SIKSPPEGFDSYIANGHTPEPDTOTLELDGQGVVYVPOGQVPVPCPEFSSSTFSOSEYLI 524
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 939 HISKLPKCKHSHVSGKGTQTPDPSS--ASITVDG--VEVPLGTGI--SSGVNDCLLYNEXIV 993
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 525 YQESQCRILYILEV 538
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 994 YDIAQVHLKYLTKL 1007

RESULT 4
S42208
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (Sarcophaga peregrina)
N.Alternate names: poly(ADP-ribose) polymerase
C.Species: Sarcophaga peregrina
C.Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 16-Jun-2000
C.Accession: S42208; S71496
R.Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurata, J.; Blochem, 220, 607-614, 1994
A.Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from Sarcophaga
A.Reference number: S42208; MUID:94170813
A.Accession: S42208
A.Molecule type: mRNA
A.Residues: 1-996 <MAS>
A.Cross-references: EMBL:DL6482; NID:g473742; PIDDN:BA03943.1; PID:g538248
A.Accession: S71496
A.Molecule type: protein
A.Residues: 170-188/721-736; 813-819; 879-885 <MAX>
C.Superfamily: NAD+ ADP-ribosyltransferase
C.Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger
F.370-507/Domain: DNA binding #status predicted <MAX>
F.508-996/Domain: auto-modification #status predicted <AMQ>

Query Match 23.6%; Score 674.5; DB 1; Length 996;
Best Local Similarity 33.6%; Pred. No. 5,1e-39;
Matches 179; Conservative 104; Mismatches 195; Indels 55; Gaps 22:

Oy 32 EEDPRSTAEALKAIPAERI-----INVDPTCLSSNPGQVY-----EDYNCTLQNTN 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 ESKSKSKSIYKSVPKSWTKLIKIDGLAVDPDSGLD--VAHVYYSRNKKEKYNVILGTTD 547
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 82 IENNNNKFFIITOLD--SNREFTCMNRGRNG-EGVQSKINHFRRLDARKDEPKKRE 138
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 548 ICKNNNSFKLIDLSDMKMRFW--FRSMGRIGTTIGGKLDNFSLNDALVQFELYLE 606
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 139 KTKNMAERDHFVSHPGKYTTLEVO--AEDEAQEAVALVVDKGRVRYVTRKVPQCSLDPAQO 197
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 607 KSGNHFENNENYKVAGRMYPTIDIDYADS-----KIDLSAEHIKSKL-PLSY-----Q 655
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 198 KIITNIFSEKEMKNTMALMDLDVKRMPGLKSLKQOIANGFLEALAEALGPTDGGSL 257
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 656 DIILKLMFVDVSKRTRMEERDLDMERKPLGKLSQKQIQSAVKVLTLEYLIQO---GGTNA 712
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 258 E--ELSSHHYVYIPINFGSOPRPINSPELLQAKKMDLLVLADIELAALQAVSDEQTV 315
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 713 KRIDATNRKRYTILPHNFGQSPPLDTEQVEQLQMDLSLIEICAVSLQATEDESKADI 772
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 316 EGVPHPLDDYDOLLSCQQLDLSGAPKRYIQTYLEQF--GSNHRPQLQ--HTRKVNQGE 372
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 773 ----NPIDKHVQLQTKLEPRLDKNSEEYILQKYKNTHAELHKLILDEYVDIIFYARQG 828
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Oy 373 EEDRFOASHKLGKRLMHGTMAVVAALITSGRLTMRP---SGRWGKGIYFASNSK 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 829 EARRRKPPEKLLNRRLLMHGSLTIFACILSHGLKIAPEAPRYCYMGKGIETADWASK 888

Qy 429 SAGVYIGMKCGAH--VGMYFLGEVALGREHHINTDNPSLSPPPGFDVSIAHGTEPDP 486

Db 889 SANVC---CTSHNSTGIMLLSEVALGDMMECTAAKYVTKLPNDK-HSCGRGRTMNP 943

Qy 487 TQDTLELDGQOYVVPQGPVPCPEFSSSTPSQSRYLLYQESQCKRLYLEVH 539

Db 944 SESIRE-DG--VEIPLGRPTINDSLKSSLL-YNFEIITYDAQVNIQYLMRN 992

RESULT 5

703656

Problem: MAD+ ADP-ribosyltransferase (EC 2.4.2.30) -- maize

C:Species: Zea mays (maize)

C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 29-Oct-1999

C:Accession: T03656

R:Babyichuk, E.; Cottrill, P.; Storozhenko, S.; Fungthong, M.; O'Farrell, M.; Van Mo

submitted to the EMBL Data Library, November 1997

A:Description: Higher plants possess two poly(ADP-ribose) polymerases.

A:Reference number: Z14991

A:Accession: T03656

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-653 <BAB>

A:Cross-references: EMBL:AJ222568; NID:e1264090; PIDN:CA10888.1; PID:e1264091

C:Genetics:

A:Gene: PARP

C:Keywords: DNA binding; glycosyltransferase; MAD; pentosyltransferase

Query Match 23.5%; Score 670.5; DB 2; Length 653;

Best Local Similarity 35.4%; Pred. No. 5.3e-39;

Matches 170; Conservative 93; Mismatches 186; Indels 31; Gaps 14;

Qy 71 EYNYCTLNQNTENNNNKFFYITQLQ--DSNRFCTQNMGRGVEGSGKINFTFLD-A 128

Db 188 ELYDATLTNQTNGDNNKFFYIIQVLES DAGSFMVYNNKRGVGRGQDKLHGSPTRDOA 247

Qy 129 KQDFKKFFREKTKNNMAERHDFHSPKYYTLEVOAEDAEAVVYKVDGPRVYTKRHVQ 188

Db 248 IYEFEGKPFNNKNNHMSDRKNKCYAKKTYWLEMDYGEFEK-----IEKG---STTDQIK 300

Qy 189 PCSLDPAFOKLTITNFSKEMERNTALMDLDVKKPKLKSQOIARGFEEALEALK 248

Db 301 EYKLETRIAQFSLICNISMKQRMVVEIGYNAEKLPGLKRAATILKGYHVLKRSIDVIS 360

Qy 249 GPTDGGQSLSELSHFYIVYIPNHGHSQPP--INSEPLDQAKKMLLVLDIELAQIQ 306

Db 361 --KADRRHLEQTLGGEYVIVPHDFGRKKRREFTIDTPQKLAKLMEVLEGEIETATKL- 417

Qy 307 AVSEQKTYVEEYRPHRDYOLKCOLDLDSGAEPEKYIQTYLEQT--GSNHRCPYLO-- 363

Db 418 ----LEDDSDDDDDLYARYKQKHCDFTPLEDSDSEYISIKRYLANTGKTHSGTYVDIV 473

Qy 364 HMKVYQGEEDRFOAHSKLGNKRLMHGCTNNAVVAALITSGLRIMPH---SGRGVSKG 419

Db 474 QTEKYSRHGEETRFQKFASTRNRMLMHGSRLSNAGLISQGLRAPPEAPYVGYMGKG 533

Qy 420 IYFASENKSAAGYVIGMKCGAHNVGYMFLGEVALGREHHINTDNPSSLSPPGFDVSIAK 479

Db 534 VYFADMFSSKSNAYCYASE--ACRSGLVLLCEVALDMMELLNADYDANNLPRGKLRSGV 591

Qy 480 GHTPEPDTQDTLELDGQOYVVPQGPVPCPEFSSSTPSQSRYLLYQESQCKRLYLEVH 539

Db 592 GQTAIPMMV-ESKVVADG--VVVPLGEPKQEPS-KRGGLLYNEYIYVNDQILMKRYVLHV 647

RESULT 6

A29725

MAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human

N:Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(RNA

C:Species: Homo sapiens (man)

C:Date: 10-Sep-1999 #sequence revision 10-Sep-1999 #text change 10-Sep-1999  
A:Accession: A29725; A28498; A39976; A26901; I38096; B33321; A33633; A61559; S14  
R:Uchida, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.;  
Biochem. Biophys. Res. Commun. 148, 617-622, 1987  
A:Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribose)  
A:Reference number: A29725; MUID:88076933  
A:Accession: A29725  
A:Molecule type: mRNA  
A:Residues: 1-69, 'Q', '71-1014 <UCH>  
A:Cross-references: GB:M8112; NID:g190166; PIDN:AAA60137.1; PID:g190167  
R:Kunusaki, T.; Ushiro, H.; Mitsunuchi, Y.; Suzuki, S.; Matsuda, M.; Matsuda, Y.; Katunuma  
J. Biol. Chem. 262, 15990-15997, 1987  
A:Title: Primary structure of human poly(ADP-ribose) synthetase as deduced from cDNA se  
A:Reference number: A28498; MUID:88058958  
A:Accession: A28498  
A:Molecule type: mRNA  
A:Residues: 1-16, 'E', '18-211, 'K', '213-236, 'R', '238-366, 'H', '369-1014 <KUR>  
A:Cross-references: GB:J03473  
R:Cherney, B.W.; McBratney, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smulson  
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987  
A:Title: cDNA sequence, protein structure, and chromosomal location of the human gene f  
A:Reference number: A39976; MUID:88066596  
A:Accession: A39976  
A:Molecule type: mRNA  
A:Residues: 1-49, 'D', '51-612, 'Q', '614-907, 'Y', '909-939, 'R', '941-979, 'T', '981-1014 <CHE>  
A:Cross-references: GB:J03030  
A:Note: the authors translated the codon ATA for residue 980 as Asn  
R:Sunuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.  
Biochem. Biophys. Res. Commun. 146, 403-409, 1987  
A:Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expression  
A:Reference number: A26901; MUID:87298455  
A:Accession: A26901  
A:Molecule type: mRNA  
A:Residues: 441-610, 'N', '612-880; 921-1014 <SUZ>  
A:Note: the sequence figure has an omission of forty residues  
R:Ogura, T.; Niyonoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.  
Biochem. Biophys. Res. Commun. 167, 701-710, 1990  
A:Title: Characterization of a putative promoter region of the human poly(ADP-ribose) p  
A:Reference number: I38096; MUID:90211250  
A:Accession: I38096  
A:Status: translation not shown  
A:Molecule type: DNA  
A:Residues: 1-40 <RES>  
A:Cross-references: EMBL:X16674; NID:9510112; PIDN:CAA34663.1; PID:g1017423  
R:Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweizer, M.  
DNA 8, 575-580, 1989  
A:Title: Human nuclear NAD(+) ADP-ribosyltransferase(polymerizing): organization of the  
A:Reference number: A33321; MUID:90091744  
A:Accession: B33321  
A:Molecule type: DNA  
A:Residues: 38-43; 93-98; 132-137; 204-209; 237-242; 276-281; 335-340; 384-389; 431-436; 512-517; 7  
A:Cross-references: GB:M29544; GB:M2953  
A:Note: the authors translated the codon GTG for residue 54 as Glu  
A:Note: these fragments represent intron-exon boundaries  
A:Accession: A33321  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 16-66; 96-121-159, 'D', '161-167 <AUD>  
A:Note: these fragments represent a zinc finger-containing DNA-binding region  
R:Grawohl, G.; Menisier de Murcia, J.; Molinete, J.; Simonin, F.; Koken, M.; Hoeljmaek  
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990  
A:Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines specific  
A:Reference number: A35635; MUID:90222155  
A:Accession: A35635  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 12-26, 'T', '28-66; 116-166 <GRA>  
R:Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Bartscher, H.J.; Hirsch  
Eur. J. Cell Biol. 44, 302-307, 1987  
A:Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.  
A:Reference number: A61559; MUID:88082900  
A:Accession: A61559  
A:Molecule type: mRNA

A:Residues:	381-420;682-710	<SCH>
R:Okoyama, Y.	Kawamoto, T.	Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Tera
Eur. J. Biochem.	194,	521-526, 1990
A:Title:	Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.	
A:Reference number:	S14010; MUID:91099327	
A:Accession:	S14010	
A>Status:	preliminary	
A:Molecule type:	DNA	
A:Residues:	1-95	<YOK>
A:Cross-references:	EMBL:X5140; NID:935286; PIDN:CA93606.1; PID:9825702	
C:Comment:	This protein can ADP-ribosylate itself as well as other proteins.	
C:Genetics:		
A:Gene:	GDB:ADPRT; PDBL	
A:Cross-references:	GDBL:119508; OMIM:173870	
A:Map position:	1q41-q42	
C:Superfamily:	NAD+ ADP-ribosyltransferase	
C:Keywords:	DNA binding; DNA repair; glycosyltransferase; NMD; nucleus; pentosyltrans	

  

Query Match	23.5%;	Score 670;	DB 1;	Length 1014;
Best Local Similarity	33.6%;	Pred. No. 1,je-38;		
Matches 186;	Conservative 101;	Mismatches 195;	Indels 72;	Gaps 23;

  

QY	7	AMAPRPKFWVOTEGPEKKGR----	OAGREDD--FRSTAEALKAIPAERILIRVDPT	58
Db	502	ALSKKSKQVVEGEEKINSEKRMLTLKGAAVDDSGLEHSAHLE-----	547	
QY	59	CPLSNPGTOYVEDYNCLNTNIENNKNKYIIOLLOD--SNEFFCWNMGVBGV-G	115	
Db	548	-----KGGKV---FSATLGLVDIVKGTNSYYKLQLLEDDEENRYW-IFRSWGVTGVI	597	
QY	116	OSKINHFEFLDAKDKEPFKEFKTKNNMAERDFHSHPGYLTIEVO-AEDEAOEAVVK	174	
Db	598	SNKLEQMSPKRDALHEHFMKLIYEKTGNAMHSK-NFTKYPRKFYLIEDYGODE--EA	654	
QY	175	VDRGFVRTVTRKVPQCSLDPAFYOKLITNISKEMFKNMTMALMDIVKMPLGKLSKO	234	
Db	655	LTVNP-GRXSKLPKP-----VDLIKMIPEVESMKRAMWEYELDIQKMPJGKLSKRO	707	
QY	235	RGFELALEALGKPTDGGSLELSHFYTIVPHNFHGSHOPPIISPELLQAKKMML	294	
Db	708	AAYSLLSYQAQAVSQSSDSOIL-DLSNRFTTLLPHDPGMKKRPLLNNABSVOAKVEM	766	
QY	295	VLADELTAQALQAVSEOEKTYEEVPHPLDRDYOLLKCOLDLSDGAREYKIQTYLEQ	353	
Db	767	NLLDIEVAVSLLRGSGSDSSKD---PIDVVEKLTADIKAVDRDSEAEIIRKYVNTH	822	
QY	354	GSNHRCPTLO--HIWKVNOGEEDRFOAHSKISGRKLTMSTNNAAVVAALITSLRI	411	
Db	823	ATTNHAIDLVIDYLFKIEREBECORYKRFKQLNHRKLRLMHGSKRTTNAGILSOGLRA	882	
QY	412	----SGRGVGKGIYFASENSKSAGYVIGMCGANH--VGYMFLGEVALGREHHINTDN	464	
Db	883	EAPYTGVMFGIGIYFADMVSKSNY-----CHTSQGDPICIGILLGEVALGMYTELKAS	936	
QY	465	SLKSPPEFDVIAIRGHETPDODTELELDGOVVVPOGQPVPCPEFSSSTFSQSXYL	524	
Db	937	HISKLPKCKHSVKGLGKTTPDPDS--ANISLDG--VDVPLGTIGT-SGCVNDTSLLYNET	991	
QY	525	YQESOCRLTYLLEY	538	
Db	992	YDIAOVNKLKYLKL	1005	

  

RESULT	7
JH0581	
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - chicken	
N:Alternate names: poly(ADP-ribose) synthase	
C:Species: Gallus gallus (chicken)	
C>Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999	
C:Accession: JH0581	
C:Title: M.E.J. Garner, J.M.; Jeltsch, J.M.; Niedergang, C.P.	
gene 102, 157-164, 1991	



C:Keywords: DNA binding; glycosyltransferase; hexosyltransferase; MAD; nucleus; pentosyl

Query Match 22.9%; Score 653; DB 2; Length 998;

Best Local Similarity 32.5%; Pred. No. 1.6e-37; Matches 177; Conservative 102; Mismatches 214; Indels 52; Gaps 20;

```

OY 10 PKPKVWVQTEGEEKKKRQAGREDFRSTALKAIPAERKIRVDPCLPSNGTQV 69
DB 480 PSSGVYAGSSSKVKEKSKNSKKMLTVGGAIDPDSGL---EDSCHVLETCG--- 533
OY 70 YEDVCTLNQNTIENNKKFYIIQLQ--DSNRFETCMNMGVAGVEY--GOSKINHFRLED 127
DB 534 -KIFSATLGLVLYTGTNTSYKQLLEHSDRSRYWYFRSGHGVATIGSKLLEMSKED 592
OY 128 AKKDEKKREKTKNNMARDFVSHPGKYTLIEVQAEDEQAEVAVKVRGVRVTKV 187
DB 593 ALEHFLNLQDTGNM--HSPNFTKPKKFPLEIDYGOE--EDVAKKLSVG--AGTKSKLA 649
OY 188 QPSLDPATOKLITNIFSKEMFKNTMALDLDVKKMPLGKLSKQOITARGFEALALEAL 247
DB 650 KP-----VQELIKLIFDVESMKKAVEFEIDLOKMPGLGKLSKQIOSAVSILSQVQAV 703
OY 248 KPTPGGGSLE---LSHFYVYIPNFGHSQPPPIPSPELLQAKKMLLVADIELA 302
DB 704 -----SELSSEARLLDSNOEYTLIPHDFGKMPPLNNLEYIOAKVOMLDNLADIEVA 757
OY 303 QAL--QAVSEOKTYEEVPHPLDRDYOLKCOLLDLSCGAPKVKQVLEOT--GSNHSC 359
DB 758 YLLLGAGADGK-----DPIDVKEKIKTIDIKVYAKDSESRITCDYKNTHADTHNA 811
OY 360 PTLQ--HIWKVQGEEDRFQASHKLGKRLKMGHTMAVVAAILTSGLRMPH---SG 413
DB 812 YDLEVLKIFPKIDREGYQYKPKFQOLHNRQLLMHGSRTTNFAGILISQGLRPAEPVVG 871
OY 414 GRVGGIYFASENSKSAGVYIGMKCAHVGTMFLGEVALGEBHNTNTNPISLKSPPGF 473
DB 872 YMFSGIYFADAVSKSANCHAMP--GSPIGLLIGEVALGNNHLEKKAASQITKL--PKGK 928
OY 474 DSVIARGHTPEPTQDTELELDGQVVPQGPVPCPEFSSSTFSGSEYLITQESQCLR 533
DB 928 HSKVGLGRAPRPS--ATVQDGS--VDVPLGKGTSA--NISDTSILYNEYIYDIAOVNLK 983
OY 534 YLLEV 538
DB 984 YLLKL 988

RESULT 10
MAD+ ADP-ribosyltransferase (EC 2.4.2.30) - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
C:Accession: S04200
R:Huppi, K.; Bialta, K.; Siwarski, D.; Kliman, D.; Cherney, B.; Smulson, M.
Nucleic Acids Res. 17, 3387-3401, 1989
A:Title: Sequence and organization of the mouse poly (ADP-ribose) polymerase gene.
A:Reference number: S04200; MUID:89263780
A:Accession: S04200
A:Molecule type: DNA
A:Residues: 1-1013 <HUP>
A:Cross-references: EMBL:X14206; NID:g49893; PIDN:CAA32421.1; PID:g49894
C:Genetics:
A:Map position: 1
C:Superfamily: MAD+ ADP-ribosyltransferase
C:Keywords: DNA binding; glycosyltransferase; MAD; nucleus; pentosyltransferase; zinc fi

```

Query Match 22.9%; Score 653; DB 1; Length 1013;

Best Local Similarity 32.3%; Pred. No. 1.7e-37; Matches 186; Conservative 102; Mismatches 211; Indels 76; Gaps 23;

OY 2 SLFLAMAPKPKPW---VOTEGPE-----KKKGROAGREEDPRSTAEALKAIP 47

DB 468 SLQDILLSAHSLSPMGAEVAYEAEVYAPRGKSAAPSKSKSGCFKEGVKS----- 518

OY 48 AKRKI-----IRVDPCLPSNFGTOYED-----YNTLNQNTIENNKKFYIIQLQ 96

DB 519 -EKRMKLTLLKGAADVPDGLSLS--AHYLEKGKVFSTLGLVIVKGTNSYKQLDLE 575

OY 97 DSNRF--FTCMNMGVAGVEY--GOSKINHFRLEDKADDFKKREKTKNNMARDFVSH 154

DB 576 DKESRYWYFRSGHGVATIGSKNLEQMPKREAEVQEPKALYEKTGNMHSK--NFTKVP 634

OY 155 GKTYLLEVO--AEDEQAEVAVKVRGVRVTKVQPCSLDPATOKLITNIFSKEMFKNT 213

DB 635 KKFYPLEIDYGOE--EAVKKTLPK--GTSKSLPKP-----VQELVGMIFVDMSKKAL 685

OY 214 ALMDLDVKKMPLGKLSKQOITARGFEALALEALGPIDGGSLELSHFYVYIPNFG 273

DB 686 VEYEDLOKMPGLGKLSRROITQAAVSTLSEVQDPVSGSSESQIL--DLSRFTLLIPHDG 744

OY 274 HSQPPPIPSPELLQAKKMLLVADIELAQAASEQKTYEVPHPPLDRDYOLKCOL 333

DB 745 MKKPLLNNADSVQAKVEKLDLDEIYAVSILRGSDSDSKD---PIDVYKLTQDI 800

OY 334 QLLDGAPEYKYIQTLYLEOT--GSNHRCPYLQ--HIWKVQGEEDRFQASHKLGKRLK 390

DB 801 KYVDRDSEAEVYIRKVKNTNATHNAVDLEVIDIFKIEREESQRYKPFROLHNRRLM 860

OY 391 HGTNNAVVAAILTSGLRMPH---SGRVGKGYFASENSKSAGVYIGMKCAH---V 443

DB 861 HOSRTTNFAGILISQGLRPAEPVGTGMFGKGYFADAVSKANSY-----CHTSGDPI 915

OY 444 GYMFLEVALGEBHINTDNPISLKSPPGFDSVIARGHTPEPTQDTELELDGQOVVPQ 503

DB 916 GLIMGEVALGNNHLEKKAASQITKL--PKGK 970

OY 504 GQVPCPEFSSSTFSGSEYLITQESQCLRILEV 538

DB 971 GTGIP--SGVNDTALYNEYIYDIAOVNLKYLKL 1004

RESULT 11

MAD+ ADP-ribosyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000

C:Accession: A47474

R:Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.

Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993

A:Title: Cloning of cDNA encoding Drosophila poly (ADP-ribose) polymerase: leucine zip

A:Reference number: A47474; MUID:93234521

A:Accession: A47474

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-994 <UCH>

A:Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:g303345; PIDN:BA02964.1; PI

A:Note: sequence extracted from NCBI Backbone (NCBIN:129703, NCBIPI:129704)

C:Genetics:

A:Gene: FLYBase:Parp

A:Cross-references: FLYBase:FBgn0010247

C:Superfamily: MAD+ ADP-ribosyltransferase

C:Keywords: DNA binding; glycosyltransferase; MAD; nucleus; pentosyltransferase

Query Match 22.4%; Score 638; DB 1; Length 994;

Best Local Similarity 32.2%; Pred. No. 1.8e-36; Matches 168; Conservative 100; Mismatches 195; Indels 58; Gaps 19;

OY 44 KAIPAEKRI-----IRVDPCTPLSSNPGTOYED-----YNTLNQNTIENNKKFYIIQL 94

DB 500 KMPKPSRTKVKVDGLAVDPDSGLEDI--AHYVDSNNKSYVGLGLDQKNNKSNYYKQL 557

OY 95 IQ--DSNRFETCMNMGVAGVEY--GOSKINHFRLEDKADDFKKREKTKNNMARDFVSH 152

Db 558 LKADKREKTYWIFRSWGRIGTINIGNSKLEEDTSESARKNEKEIYADKGTNEYBQDRNFVK 617  
Qy 153 HPGKTYLIEVQAEDEAOEAVYKVDGRPVRTVTRKQPCSLDPATOKLTITNFSKEMKNT 212  
Db 618 RGRMRPIELQYDD--QKLVKHE-----SHEFTSKLEISVQNLKILFIPIIDSMNK 667  
Qy 213 MALMDIDVKKMPLGKLSKQOIALRGFEALBEPALBEPDGGOSLELSHFYTVIPHNE 272  
Db 668 LMFEHIDMDKMPGLKLSAQIOGAYRVKVEIYVLECGSNFATK-LIDATNREFTLLPHNE 726  
Qy 273 GHSQPPINSPELLQAKKMLVLADIELAOIAVSEOKYEEVPHPLDRDYQLKCO 332  
Db 727 GYOLPLTIFHOQIEDLRMLDLSLEIYVAYSI-----IKSEYSDACNPLDNHYAOKTQ 782  
Qy 333 LQLDGSGAEYKIQYLYGOT--GSNRCPTIQ--HIMKXNOEGEEDRFQAHSLGKRL 389  
Db 783 LVALDENSEEFTLSQYVNTASTHSTKSYDLKIVDFKYSKQGEARFEPFKKLNKRL 842  
Qy 390 MHGTNNAVVAALITSLRI---MPHSGRVGKGIYFASENSKAGYVIGMKGAHNVG 445  
Db 843 MHGSRITNFGILSHGLRIAPRAPPTGYMFGKGIYFADNVKSAHYCCTSQ--QNSTGL 900  
Qy 446 MFLGEVALG-----REHHIN--TDNPSLKSPPPGFDSVIARGHTEPPTQDTLELDQ 497  
Db 901 MLTSEVALDMMECTSAAKYNKLSNNK-----HSCFGRGRTMPDPTK-SYIRSDG- 949  
Qy 498 QVYVPGQVPYPCPEFSSSTFSSQSEVLYQESQCRLEYLEY 538  
Db 950 -VEIPYGTI-TDEHLKSLLYNEYIVYDVAVQVNIQYLFPM 988

## RESULT 12

T51353  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) [imported] - Arabidopsis thaliana  
M:Alternate names: poly(ADP-ribose) polymerase  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: T51353  
R:Doucet-chabeaud, G.; Kazmaier, M.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z23379  
A:Accession: T51353  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-983 <DOU>  
A:Cross-references: EMBL:AJ131705; PIDN:CAA10482.1  
A:Experimental source: cultivar Landsberg erecta  
C:Genetics:  
A:Gene: parp-1  
C:Function:  
A:Description: ADP-ribose polymer synthesis  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match 21.0%; Score 598.5; DB 2; Length 983;  
Best Local Similarity 31.1%; Pred. No. 1e-33;  
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;  
Qy 24 KKGROAGREDDPFRSTAALKAIPAERKRIIRVDPCTPLSSNPGTO---VYED---YNC 75  
Db 472 KQQRKLPFKYKIEDTSESLVYKVGK-----SAVHEASGLQCHILLEDGNSIYNT 524  
Qy 76 TLNQTNIENNKKFYIIQLDLSNRFETGW--NRMGVGV--EYGOSKINHFRLEDAKDD 131  
Db 525 TLSMSDLSTGINSYLLQIIQ--DKGSDCYVFRKMGVGNENKIGKRVEMSK-SDAVHE 582  
Qy 132 FEKKFREKTKN--NMAEDDHVSHGKTYLLEVOAEDBAOEAVVYVNDGVPRTYVKRQ 188  
Db 583 FRRLFEKTKGNWSEWQKTNFOKOPGKFLPD-----IDGVKKQVAKK-E-628  
Qy 189 P-----CSLDPATOKLTITNFSKEMKNTMALMDLDVKKMPLGKLSKQOJARGEFALEALE 244

Db 629 PFOSSNLAPSLIELMKMLFDVETYSAMEFEINNSEMPGLKLSKHNIQKGFALTEIQ 688  
Qy 245 EAL-----KGPDOGOSLELSHFYTVIPHNFHSOPPPINSPELLQAKMDLVADI 300  
Db 689 RLTTESDPOPTMKESLLVDNSNRFETWIP---SIPHHIIRDEDDKSKYKMLEADIE 744  
Qy 301 LAQALQAVSEOEKTEVEVPHPLDRDYQLKCOLQLDGAPKRYQTLQEQTSNHRP 360  
Db 745 IASRI--VEEDVDSTES-----LDDKRYKLHCDISPLPHSEDRLEKTLANT---HNP 794  
Qy 361 T-----LQHTWKXNOEGEEDRFQAH-SKLGKRLIMHGTNNAVVAALITSLRI---- 408  
Db 795 THTWMSLELEVVALREGESEFQKAPHREKLGKMLMHGSRITNFGILNGLRIAPPE 854  
Qy 409 MPHSGRVGKGIYFASENSKAGYVIGMKGAHNVGTYMFLGVALGREHINTDNLSKS 468  
Db 855 APATGYMFGKGIYFADLVSKSAQY--YTCCKNPVGLMLTSEVALGEIHEL-TRAKYMK 911  
Qy 469 PRPGFDSVIARGHTEPPTQDTLELDQOQVYVPGQVPYPCPEFSSSTFSSQSEVLYQES 528  
Db 912 PRPGKSTGTGLKQKVP---QDSEFAKMRGDYVPCGKPV-SKRYKASLEMYNEYIYDPA 967  
Qy 529 QCRRLRYLEY 538  
Db 968 QVYKQFLKLV 977

## RESULT 13

C84719  
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C84719  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Molnar, R.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon,  
euss, D.; Niernman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: C84719  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1009 <STO>  
A:Cross-references: GB:AE002093; NID:g4433827; PIDN:AAD20677.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g31320  
A:Map position: 2  
C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 21.0%; Score 598.5; DB 2; Length 1009;  
Best Local Similarity 31.1%; Pred. No. 1e-33;  
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;  
Qy 24 KKGROAGREDDPFRSTAALKAIPAERKRIIRVDPCTPLSSNPGTO---VYED---YNC 75  
Db 498 KQQRKLPFKYKIEDTSESLVYKVGK-----SAVHEASGLQCHILLEDGNSIYNT 550  
Qy 76 TLNQTNIENNKKFYIIQLDLSNRFETGW--NRMGVGV--EYGOSKINHFRLEDAKDD 131  
Db 551 TLSMSDLSTGINSYLLQIIQ--DKGSDCYVFRKMGVGNENKIGKRVEMSK-SDAVHE 608  
Qy 132 FEKKFREKTKN--NMAEDDHVSHGKTYLLEVOAEDBAOEAVVYVNDGVPRTYVKRQ 188  
Db 609 FRRLFEKTKGNWSEWQKTNFOKOPGKFLPD-----IDGVKKQVAKK-E-654  
Qy 189 P-----CSLDPATOKLTITNFSKEMKNTMALMDLDVKKMPLGKLSKQOJARGEFALEALE 244  
Db 655 PFOSSNLAPSLIELMKMLFDVETYSAMEFEINNSEMPGLKLSKHNIQKGFALTEIQ 714  
Qy 245 EAL-----KGPDOGOSLELSHFYTVIPHNFHSOPPPINSPELLQAKMDLVADI 300  
Db 715 RLTTESDPOPTMKESLLVDNSNRFETWIP---SIPHHIIRDEDDKSKYKMLEADIE 770



```
Db      648  C T L P A G F Q S V Q G V G R Q C P R E I G S Y K N P D G Y T V P L G I T Y M Q L G K Q N V ----- 694
Oy      514  S T F S Q S E Y L I Y Q S Q C R L R Y L L E V H L 540
Db      695  D Y H L L Y N E F I Y V D Y D I O L K Y L V R Y K M 721
```

Search completed: August 29, 2002, 07:59:33  
Job time: 280 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:57:16 ; Search time 124.84 Seconds  
(without alignments)  
474.226 Million cell updates/sec

Title: US-09-701-586b-8  
Perfect score: 2813  
Sequence: 1 MAPKKKASVQTEGSKKQROG.....EYLYIKESQRLRYLLEIHL 533

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2813	100.0	533	21 AAY51177	Murine PARP1 (Long
2	2776.5	98.7	528	21 AAY51178	Murine PARP1 (short
3	2241	79.7	533	21 AAY51175	Human brain PARP3
4	2241	79.7	540	21 AAY51176	Human uterus type
5	2229	79.2	533	22 AAU29021	Human PARP-3 prote
6	717	25.5	653	21 AAY68834	A poly(ADP-ribose)
7	705	25.1	1013	17 AAR99642	Poly(ADP-ribose) p
8	702.5	25.0	522	22 AAU29022	Mouse PARP-2 prote
9	700	24.9	1014	21 AAY58043	Human poly (ADP-ri
10	700	24.9	1014	22 AAU29019	Human PARP-1 prote
11	700	24.9	1014	22 AAB66296	Human tankyrase2 r

12	699	24.8	1014	20 AAY33699	Human poly(ADP-rib
13	698	24.8	521	22 AAB60693	Human poly(ADP-rib
14	698	24.8	570	21 AAY51174	Human brain PARP2
15	698	24.8	570	22 AAB11480	Human brain poly-A
16	696.5	24.8	534	21 AAB42909	Human ORFX ORF2673
17	696.5	24.8	534	22 AAU29023	Human PARP-2 prote
18	696.5	24.8	583	22 AAB47029	hPARP2. Homo sapi
19	696	24.7	1014	21 AAY49939	Human nuclear NAD+
20	695.5	24.7	534	22 AAU29020	Human PARP-2 prote
21	690	24.5	637	21 AAY68835	The poly(ADP-ribos
22	646.5	23.0	531	22 AAB93513	Human protein sequ
23	619	22.0	557	22 ABB66431	Drosophila melanog
24	599.5	21.3	1063	22 AAB47032	Fusion protein PAR
25	584.5	20.8	969	21 AAY68833	A poly(ADP-ribose)
26	584.5	20.8	980	21 AAY68839	A poly(ADP-ribose)
27	491	17.5	982	20 AAY28464	Maize poly ADP-rib
28	456	16.2	379	22 AAU21687	Novel human neopla
29	435	15.5	360	22 AAB47030	N-terminal fragmen
30	398	14.1	1010	21 AAY68840	Fusion protein of
31	378	13.4	294	22 AAU20129	Human DNA repair a
32	378	13.4	294	22 AAU21810	Novel human neopla
33	364.5	13.0	287	22 AAB47031	C-terminal fragmen
34	275	9.8	1730	22 AAU3242	Novel human secret
35	270.5	9.6	1724	21 AAY54373	cDNA sequence enco
36	270.5	9.6	1724	22 AAB51022	Human minor vault
37	261	9.3	1099	22 AAB66301	Human tankyrase2 e
38	246.5	8.8	227	22 AAU20130	Human DNA repair a
39	246.5	8.8	227	22 AAU21811	Novel human neopla
40	203.5	7.2	190	22 AAU21686	Novel human neopla
41	156.5	5.6	1181	22 ABB60894	Drosophila melanog
42	156.5	5.6	1181	22 AAB66297	Drosophila tankyra
43	146	5.2	1327	21 AAB27212	Human tankyrase I
44	146	5.2	1327	21 AAY44402	Human tankyrase.
45	146	5.2	1327	22 AAB66279	Human tankyrase1 S

ALIGNMENTS

RESULT 1  
AAY51177  
ID AAY51177 standard; Protein: 533 AA.  
XX  
AC AAY51177;  
XX  
DT 31-MAR-2000 (first entry)  
XX  
DE Murine PARP1 (long) homologue protein.  
XX  
PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP1.  
XX  
OS Mus sp.  
XX  
PN WO9964572-A2.  
XX  
PD 16-DEC-1999.  
XX  
PF 04-JUN-1999; 99WO-EP03889.  
XX  
PR 05-JUN-1999; 98DE-1025213.  
PR 01-MAR-1999; 99DE-1008837.  
PA (BADI ) BASF AG.  
XX  
Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX  
WPI: 2000-087218/07.  
XX  
DR N-PSDB; AAZ44290.  
XX  
Novel genes and proteins, antibodies and binding partners useful in  
PT diagnosis and therapy of energy deficiency associated disease

PT conditions -  
 PS Claim 4; Page 67-69; 96pp; German.  
 XX  
 CC This invention describes novel human and murine poly(ADP-ribose)  
 CC polymerase (PARP) homologues, which are characterised by an amino acid  
 CC sequence motif, of general formula NAD<sup>+</sup>-binding site and no zinc finger  
 CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>C (I). The nucleic acid  
 CC sequences, PARP homologues and antibodies are useful for analytic  
 CC detection of PARP homologues and for identifying PARP effectors or  
 CC binding partners, as well as for determining their effectiveness.  
 CC PARP-binding partners are useful for the diagnosis or therapy of a  
 CC disease condition, which is the result of a PARP protein, especially an  
 CC energy deficiency, which may comprise tissue damage from cell death  
 CC following necrosis or apoptosis. The disease condition may be chosen  
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
 CC in particular neurotoxic disturbances, etc. This sequence represents the  
 CC murine PARP1 protein used in the method of the invention.  
 XX  
 SQ Sequence 533 AA;  
 Query Match 100.0%; Score 2813; DB 21; Length 533;  
 Best Local Similarity 100.0%; Pred. No. 3e-243;  
 Matches 533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MAPKKASVOTEGSKKORQTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIOV 60  
 DB 1 mapkrkasvqtgskkqrqtgdeedsfrstaealraapadnrvirvdpscfgrnpgiqr 60  
 QY 61 HEDYDCTLNQTNIGNNNKFFYIIQLLEGGSRFFCWNRWGRVGEVQSGKMNHFTCLEDAKK 120  
 DB 61 hedydctlnqtnignnnkfyyiqlleegsrffcwrrwgrvgevqsgkmnhftcleadakk 120  
 QY 121 DFKKKFWEKTKNWEERDRFVAQPNKYTLIEVGEAESQAVKALSPQVDSGPRVTVK 180  
 DB 121 dfkkkfwektnkweerdrrfvagpnktyllievgeaesqavvkalspqvdsgrvtrtvk 180  
 QY 181 PCSLDPATONLTNIFSKEMFNKMTLMNLDVKMKPLGKLTQKQIARGFEALEEAMK 240  
 DB 181 pcsldpatqnltnifskemfknamtlnmldvkkmplgkltkqkqargfealeeakm 240  
 QY 241 NPTGDGQSEELSCCYTVIPIHNFGRSRPPPIINSPDVLOAKKMDMLLVADIETLAQTLOAA 300  
 DB 241 nptgdgqseelsscctvtyviphnfgsrppppinspdvloakmdmlvladielaqtlqaa 300  
 QY 301 PGESEKVEEVPHPDLRDYQLLRQQLDLSGESEYKAIQTYLKQTNNSRCPNLRHVWK 360  
 DB 301 pgeeeekveevphpldrdyqlrrcqqlldsgeseykaiqtylkqtnnsrncplnrhvwk 360  
 QY 361 VNREGGDRFOAHSKLGRRLLWHTNVAATLTSLGRIMPHSGRGVKGIFYASENS 420  
 DB 361 vnreggdrrfoahsklgrrllwhtnvaatlttsglrmpshsgrgvgkyfyasens 420  
 QY 421 KSAGYVTTMHCQGQVGMFLGEVALGKEHHTIDDPKSPPPGDSVIARGQTEPDPA 480  
 DB 421 ksagyvttmhcgqgvgmflgevalgkehhtiddpkspppgdsviargqtepdpa 480  
 QY 481 QDIELEDGQPVVPQGPVQCPFSKSSFSQSEYLYIKESQCRRLVLEIHL 533  
 DB 481 qdieleldgqpvvpqgpvpqcpfskssfsqseyllyikesqcrlylleihl 533

RESULT 2

AAV51178  
 ID AAY51178 standard; Protein; 528 AA.  
 XX  
 AC AAY51178;  
 XX  
 DT 31-MAR-2000 (first entry)  
 XX Murine PARP1 (short) homologue protein.  
 XX  
 DE

KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
 KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
 KW ischemic tissue damage; PARP1.  
 XX Mus sp.  
 OS  
 XX WO9964572-A2.  
 PN  
 XX 16-DEC-1999.  
 PD  
 XX 04-JUN-1999; 99WO-EP03889.  
 PF  
 XX 05-JUN-1998; 98DE-1025213.  
 PR  
 XX 01-MAR-1999; 99DE-1008837.  
 PP  
 XX (BADI ) BASF AG.  
 PA  
 XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
 PI WPI; 2000-087218/07.  
 DR N-PSDB; AA244291.  
 DR  
 XX Novel genes and proteins, antibodies and binding partners useful in  
 PT diagnosis and therapy of energy deficiency associated disease  
 PT conditions -  
 PT  
 XX Claim 4; Page 71-73; 96pp; German.  
 PS  
 XX This invention describes novel human and murine poly(ADP-ribose)  
 CC polymerase (PARP) homologues, which are characterised by an amino acid  
 CC sequence motif, of general formula NAD<sup>+</sup>-binding site and no zinc finger  
 CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>C (I). The nucleic acid  
 CC sequences, PARP homologues and antibodies are useful for analytic  
 CC detection of PARP homologues and for identifying PARP effectors or  
 CC binding partners, as well as for determining their effectiveness.  
 CC PARP-binding partners are useful for the diagnosis or therapy of a  
 CC disease condition, which is the result of a PARP protein, especially an  
 CC energy deficiency, which may comprise tissue damage from cell death  
 CC following necrosis or apoptosis. The disease condition may be chosen  
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
 CC in particular neurotoxic disturbances, etc. This sequence represents the  
 CC murine PARP1 protein used in the method of the invention.  
 XX  
 SQ Sequence 528 AA;  
 Query Match 98.7%; Score 2776.5; DB 21; Length 528;  
 Best Local Similarity 99.1%; Pred. No. 5.5e-240;  
 Matches 528; Conservative 0; Mismatches 0; Indels 5; Gaps 1;  
 QY 1 MAPKKASVOTEGSKKORQTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIOV 60  
 DB 1 mapkrkasvqtgskkqrqtgdeedsfrstaealraapadnrvirvdpscfgrnpgiqr 60  
 QY 61 HEDYDCTLNQTNIGNNNKFFYIIQLLEGGSRFFCWNRWGRVGEVQSGKMNHFTCLEDAKK 120  
 DB 61 hedydctlnqtnignnnkfyyiqlleegsrffcwrrwgrvgevqsgkmnhftcleadakk 120  
 QY 121 DFKKKFWEKTKNWEERDRFVAQPNKYTLIEVGEAESQAVKALSPQVDSGPRVTVK 180  
 DB 121 dfkkkfwektnkweerdrrfvagpnktyllievgeaesqavvkalspqvdsgrvtrtvk 180  
 QY 181 PCSLDPATONLTNIFSKEMFNKMTLMNLDVKMKPLGKLTQKQIARGFEALEEAMK 240  
 DB 181 pcsldpatqnltnifskemfknamtlnmldvkkmplgkltkqkqargfealeeakm 240  
 QY 241 NPTGDGQSEELSCCYTVIPIHNFGRSRPPPIINSPDVLOAKKMDMLLVADIETLAQTLOAA 300  
 DB 241 nptgdgqseelsscctvtyviphnfgsrppppinspdvloakmdmlvladielaqtlqaa 300  
 QY 301 PGESEKVEEVPHPDLRDYQLLRQQLDLSGESEYKAIQTYLKQTNNSRCPNLRHVWK 360  
 DB 301 pgeeeekveevphpldrdyqlrrcqqlldsgeseykaiqtylkqtnnsrncplnrhvwk 360  
 QY 361 VNREGGDRFOAHSKLGRRLLWHTNVAATLTSLGRIMPHSGRGVKGIFYASENS 420  
 DB 361 vnreggdrrfoahsklgrrllwhtnvaatlttsglrmpshsgrgvgkyfyasens 420  
 QY 421 KSAGYVTTMHCQGQVGMFLGEVALGKEHHTIDDPKSPPPGDSVIARGQTEPDPA 480  
 DB 421 ksagyvttmhcgqgvgmflgevalgkehhtiddpkspppgdsviargqtepdpa 480  
 QY 481 QDIELEDGQPVVPQGPVQCPFSKSSFSQSEYLYIKESQCRRLVLEIHL 533  
 DB 481 qdieleldgqpvvpqgpvpqcpfskssfsqseyllyikesqcrlylleihl 533

QY 361 VNREGGDRFOAHSKLGNNRLLWHGTNVAVVAAIITSLGRIMPHSGGVRGKGIYFASNS 420  
 DB 356 vnreggdrfahsklgnnrllwhgtnvavvaaltsglrmpshsggvrvgkgyifasns 415  
 QY 421 KSAGYVVTMHCGGHQGVGMFLGEVALGKEHHITIDDPKSLKSPPPGFDVSIARGOTEPDPA 480  
 DB 416 ksagvvtmhcgghqgvgygmflgevalgkehnhitiddpsalkpppgfdsviargotepdpa 475  
 QY 481 ODIELELDGQPVVVPVPGPPVQPCPFKSSFSQSEYLIYKESQCRRLRYLLEIHL 533  
 DB 476 qdieleldgqpvvvpvpgppvqpcpfkssfsqseyliykesqcrlylleihl 528

RESULT 3  
 AAY51175  
 ID AAY51175 standard; Protein: 533 AA.  
 AC AAY51175;  
 DT 31-MAR-2000 (first entry)  
 DE Human brain PARP3 protein.  
 KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
 KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
 KW ischemic tissue damage; PARP3.  
 OS Homo sapiens.  
 XX  
 PN WO9964572-A2.  
 XX  
 PD 16-DEC-1999.  
 XX  
 PF 04-JUN-1999; 99WO-EP03889.  
 PR 05-JUN-1998; 98DE-1025213.  
 PR 01-MAR-1999; 99DE-1008837.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
 XX  
 DR WPI; 2000-087218/07.  
 DR N-PSDB; AAZ44288.  
 XX  
 PT Novel genes and proteins, antibodies and binding partners useful in  
 PT diagnosis and therapy of energy deficiency associated disease  
 PT conditions -  
 PS Claim 4; Page 57-59; 96pp; German.  
 XX  
 CC This invention describes novel human and murine poly(ADP-ribose)  
 CC polymerase (PARP) homologues, which are characterised by an amino acid  
 CC sequence motif with a functional NAD<sup>+</sup>-binding site and no zinc finger  
 CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>2C (1). The nucleic acid  
 CC sequences, PARP homologues and antibodies are useful for analytic  
 CC detection of PARP homologues and for identifying PARP effectors or  
 CC binding partners, as well as for determining their effectiveness.  
 CC PARP-binding partners are useful for the diagnosis or therapy of a  
 CC disease condition, which is the result of a PARP protein, especially an  
 CC energy deficiency, which may comprise tissue damage from cell death  
 CC following necrosis or apoptosis. The disease condition may be chosen  
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
 CC in particular neurotoxic disturbances, etc. This sequence represents the  
 CC human PARP3 protein used in the method of the invention.  
 XX  
 SQ Sequence 533 AA;

Query Match 79.7%; Score 2241; DB 21; Length 533;  
 Best Local Similarity 80.1%; Pred. No. 5.8e-192;  
 Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

QY 1 MAPRKASVOTEG--SKKROGTREEDSFRSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58  
 DB 1 maprkspvwqteggpkxkgrqagreedpfrstaealkaipaekrlirvdpccpssnpgt 60  
 QY 59 QVHEDYDCTLQNTIGNNNNKFIYIIQLEEGSRFF-CWNRWGRVGEVQSGKMNHFTCLIED 117  
 DB 61 qvyedynctlnqtniennnnkfyiiqlqdsnrfftcwnrwrgrvgevgqskinhfrl 120  
 QY 118 AKKDFKKFKWEKTKNKEERDRFVAQPNKYTLIEVQGEAESQEAQVAVKALSQVDSGPPVRT 177  
 DB 121 akkdfekfkfkrktnnwaerdhfvshpgkytlievqaedeaqvavk-----vdrgrpvrt 175  
 QY 178 V---VKPCSLDPATQNLITNIFSKEMFKNAMTLMNLDVKKMPLGKLTQKQIARGFEALEA 234  
 DB 176 vtkrvqpcslpatcqkltitnifskemfntmalmldvkkmplgklskqqlargfealea 235  
 QY 235 LEEAMKNPTGDGSLSEELSSCFYTVIPHNFGRSRPPPIPSDVLOAKKMDLLVLADIELA 294  
 DB 236 leealkgptdggqslseelssshfytviphnfghsqpppkinspellqakkmllvliadiela 295  
 QY 295 QTLQAAPGEEEEKVEEVPPLDRYQQLLRQQLQLLDGSESEYKAIQYLYKOTGNSYRCPN 354  
 DB 296 qalqav-segektveevphpldrdyqlkqqlldsgapeykvityledtgnhrct 354  
 QY 355 LRHVYKNRREGGDRFOAHSKLGNNRLLWHGTNVAVVAAIITSLGRIMPHSGGVRGKGIY 414  
 DB 355 lqhiwkvndgegedrfahsklgnnrllwhgtnvavvaaltsglrmpshsggvrvgkgyi 414  
 QY 415 FASENSKSAGYVVTMHCGGHQGVGMFLGEVALGKEHHITIDDPKSLKSPPPGFDVSIARGQ 474  
 DB 415 fasensksagyyvmkgcghhgygmflgevalgrehintdnpslkappgfdsviargh 474  
 QY 475 TEPDPAODIELELDGQPVVVPVPGPPVQPCPFKSSFSQSEYLIYKESQCRRLRYLLEIHL 533  
 DB 475 tepdptdteledgqpvvvpvpgppvqpcpfkssfsqseyliykesqcrlylleihl 533

RESULT 4  
 AAY51176  
 ID AAY51176 standard; Protein: 540 AA.  
 AC AAY51176;  
 DT 31-MAR-2000 (first entry)  
 DE Human uterus type 2 PARP3 protein.  
 KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
 KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
 KW ischemic tissue damage; PARP3.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9964572-A2.  
 XX  
 PD 16-DEC-1999.  
 XX  
 PF 04-JUN-1999; 99WO-EP03889.  
 PR 05-JUN-1998; 98DE-1025213.  
 PR 01-MAR-1999; 99DE-1008837.  
 XX  
 PA (BADI ) BASF AG.  
 XX  
 PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
 XX  
 DR WPI; 2000-087218/07.  
 DR N-PSDB; AAZ44289.  
 XX  
 PT Novel genes and proteins, antibodies and binding partners useful in  
 PT diagnosis and therapy of energy deficiency associated disease  
 PT conditions -

XX Claim 4; Page 62-64; 96pp; German.

XX This invention describes novel human and murine poly(ADP-ribose)

CC polymerase (PARP) homologues, which are characterised by an amino acid

CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger

CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MHX<sub>2</sub>C (I). The nucleic acid

CC sequences, PARP homologues and antibodies are useful for analytic

CC detection of PARP homologues and for identifying PARP effectors or

CC binding partners, as well as for determining their effectiveness.

CC PARP-binding partners are useful for the diagnosis or therapy of a

CC disease condition, which is the result of a PARP protein, especially an

CC energy deficiency, which may comprise tissue damage from cell death

CC following necrosis or apoptosis. The disease condition may be chosen

CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,

CC in particular neurotoxic disturbances, etc. This sequence represents the

CC human PARP3 protein used in the method of the invention.

XX

SQ Sequence 540 AA;

Query Match 79.7%; Score 2241; DB 21; Length 540;

Best Local Similarity 80.1%; Pred. No. 6e-192;

Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

QY 1 MAPKRKASVQTEG--SKKORQGTEDSFSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58

DB 8 mapkpwvqtgepekkkgqagreedfrstaealkaipaekriirvdtcptlsnpgt 67

QY 59 QVHEDYDCTLNQTNIGNNNKFFYIIQLLEGSRFF-CWNRMGRVGEVGSKMNHFTCLEL 117

DB 68 qvyedyncntlnqtniennnnkffylqlldgsnfftcwnrgrvgrvgqskinhfrled 127

QY 118 AKKDFKKFKFEKTKNWEERDRFVAOPNKYTLLEVOGEAESQAVVKALSPQVDSGPVRT 177

DB 128 akkdfekfkfektknwaerdhrvshpgkytlievgaedaevavk-----vdrapvrt 182

QY 178 V---VKPCSLDPATQNLITNIFSKEMFKNMTLMNLDVKKMPLGLTKTKQIARGFFALEA 234

DB 183 vtrkvqpsldpatqklitnifskemfkmtalmdldvkkmpgkiskqargfealea 242

QY 235 LEAMKNPTGDGQSLSELSCTFYVTPHNFGRSRPPINSPOVLQAKKMDLLVLADIELA 294

DB 243 lealkgptdggsgleelsshfytviphnfghsqppinspellqakkdmlvladiela 302

QY 295 QTLQAAPGEEERKEVEVPHPDLDRDYOLLRCQLQLDSEGESEYKAIOTYLTGKTSYRCPN 354

DB 303 qalqav-seqektveevphpldrdyqlkqqlldsgapeykvigtyleqtsghrcpt 361

QY 355 LRHWKYNREGEDRFQAHSKLGNRRLLNHTGNTVAVVAAILTSGLRIMPHSGRGVKGIIY 414

DB 362 lqhiwkvngqeedrfaqhsklgnrkllwhgtmavvaailtsglrimphsggrvgkgy 421

QY 415 FASENSKAGVYVTHMCGGHGVGMELGEVALGKEHHITIDDSLKSPPPGDSVTARGQ 474

DB 422 fasensksagvylgmkgahvdmflgevalgrhnhintdpslksppgfdsvlargh 481

QY 475 TEPDPAQIDLELDGPPVVPQGPVQCPSEKSSFSQSSELYLYKESQRLRYLLRIHL 533

DB 482 tepdptqdtelldgqvvvpgqgpcpsefsstfsqseyllyqesqcrilyllehl 540

RESULT 5

AAU29021

ID AAU29021 standard; Protein; 533 AA.

XX

AC AAU29021;

XX

DT 18-DEC-2001 (first entry)

XX

DE Human PARP-3 protein.

XX

KW Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;

KW cytostatic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;

KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;

KW oxidative stress; neurological disorder; parkinsonism; apoptosis;

KW meningitis-associated intracranial complication; ischaemia;

KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.

XX Homo sapiens.

OS WO200164955-A1.

PN 07-SEP-2001.

XX

PD 01-MAR-2001; 2001WO-US06572.

PF 02-MAR-2000; 2000US-0517467.

XX (ISIS-) ISIS PHARM INC.

PA Popoff I, Cowsert LM;

PI WPI: 2001-602570/68.

XX N-PSDB; AAS45590.

DR

DR Antisense compound useful for treating hyperproliferative, neurological, inflammatory and autoimmune disorders and diabetes

XX inhibits human PARP -

PT

PT

PS Example 13; Page 105-107; 168pp; English.

XX

CC The invention relates to antisense oligonucleotides targeted to human

CC PARP nucleic acid and inhibiting expression of human PARP. PARP

CC (Poly (ADP-ribose) polymerase plays an important role in chromatin

CC decondensation, DNA replication, DNA repair, gene expression, malignant

CC transformation, cellular differentiation and apoptosis. The antisense

CC oligonucleotide inhibitors are useful for inhibiting the expression of

CC PARP in human cells or tissues. They are also useful for treating a

CC human with a disease associated with PARP especially hyperproliferative

CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,

CC neurological (e.g parkinsonism, meningitis-associated intracranial

CC complications and ischaemia), inflammatory and autoimmune disorders (e.g

CC arthritis) and diabetes. The present sequence is a PARP protein,

CC the cDNA encoding which was used to design the antisense

CC oligonucleotides.

XX

SQ Sequence 533 AA;

Query Match 79.2%; Score 2229; DB 22; Length 533;

Best Local Similarity 79.8%; Pred. No. 7e-191;

Matches 430; Conservative 37; Mismatches 60; Indels 12; Gaps 5;

QY 1 MAPKRKASVQTEG--SKKORQGTEDSFSTAEALRAAPADNRVIRVDPSCPFSRNPGI 58

DB 1 mapkpwvqtgepekkkgqagreedfrstaealkaipaekriirvdtcptlsnpgt 60

QY 59 QVHEDYDCTLNQTNIGNNNKFFYIIQLLEGSRFF-CWNRMGRVGEVGSKMNHFTCLEL 117

DB 61 qvyedyncntlnqtniennnnkffylqlldgsnfftcwnrgrvgrvgqskinhfrled 120

QY 118 AKKDFKKFKFEKTKNWEERDRFVAOPNKYTLLEVOGEAESQAVVKALSPQVDSGPVRT 177

DB 121 akkdfekfkfektknwaerdhrvshpgkytlievgaedaevavk-----vdrapvrt 175

QY 178 V---VKPCSLDPATQNLITNIFSKEMFKNMTLMNLDVKKMPLGLTKTKQIARGFFALEA 234

DB 176 vtrkvqpsldpatqklitnifskemfkmtalmdldvkkmpgkiskqargfealea 235

QY 235 LEAMKNPTGDGQSLSELSCTFYVTPHNFGRSRPPINSPOVLQAKKMDLLVLADIELA 294

DB 236 lealkgptdggsgleelsshfytviphnfghsqppinspellqakkdmlvladiela 295

QY 295 QTLQAAPGEEERKEVEVPHPDLDRDYOLLRCQLQLDSEGESEYKAIOTYLTGKTSYRCPN 354

Db 296 qalqav-seqektveevphldrdyqlkcklqllldsgapeykvqityleqtgshnropt 354  
QY 355 LRHVVKVNRREGSGDFQAHSKLGNRLWLHGTVNAVVAAILTSGLRIMPHSGRGVKGIIY 414  
Db 355 lqhwkvvqgeedrfqahsklgnrkllwhgtnmavvaailtsglrimphsgrgvkgiiy 414  
QY 415 FASENSKASAGYVTTMHCQGHQGVYMFGLGEVALGKEHHITIDDPSSLKSPPPGDFSDVIARGQ 474  
Db 415 fasensksagyyigmkcgahhvymflgevalgrehhintdnpslkspppgfdsviargh 474  
QY 475 TPDPADQIDELDGQVVPVPGPPVQCPFSKSSFSQSEYLIYKESQCRRLYLLEIHL 533  
Db 475 tepdptqdtelldgqvvvpgqpvpcefasstfsqseyliyqesqrilylievhl 533

RESULT 6  
AA68834  
ID AAY68834 standard; Protein; 653 AA.  
AC AAY68834;  
XX  
XX 16-MAY-2000 (first entry)  
XX  
DE A poly(ADP-ribose) polymerase NAP protein of Zea mays.  
XX  
XX NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;  
KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;  
KW pest; drought; heat; fungi; nematode; seed-shatter.  
XX  
OS Zea mays.  
FH  
FH Key Location/Qualifiers  
FT Misc-difference 1..159  
FT /note= "these residues are specifically claimed in  
FT claim 18"

XX  
XX WO200004173-A1.  
XX  
XX 27-JAN-2000.  
XX  
XX 12-JUL-1999; 99WO-EP04940.  
XX  
XX 17-JUL-1998; 98US-0118276.  
XX  
XX (PLBZ ) PLANT GENETIC SYSTEMS NV.  
XX  
XX Babilychuk E, Kushnir S, De Block M;  
XX  
XX WPI; 2000-182436/16.  
XX  
XX N-PSDB; AA260616.  
XX  
XX Modulating cell death, growth and stress resistance in eukaryotes,  
XX specifically plants, used, e.g. to impart fungus or nematode resistance  
XX  
XX Claim 18; Page 92-95; 126pp; English.  
XX  
XX The present sequence represents the NAP protein of Zea mays. This  
XX protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as  
XX poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed  
XX cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide  
XX sequences can be used for modulation of programmed cell death in  
XX eukaryotic cells. The method is used, specifically in plants, to induce,  
XX or protect against, programmed cell death, depending on the extent to  
XX which PARP activity is reduced. Reducing expression of endogenous NAP  
XX class PARP only is also used to modulate programmed cell death, to  
XX increase growth rate and to produce plant cells that are more tolerant  
XX of stress (cold, chemical treatments, pathogens, pests, drought, heat,  
XX etc., or during transformation). Particular applications are generation  
XX of plants that are resistant to fungi or nematodes; are male or female  
XX sterile; or have better seed-shatter properties. The methods are also  
XX used to improve growth of transformed plant cells (and derived calli or  
XX complete plants).

XX  
SQ Sequence 653 AA;  
  
Query Match 25.5%; Score 717; DB 21; Length 653;  
Best Local Similarity 34.9%; Pred. No. 3.7e-55;  
Matches 190; Conservative 91; Mismatches 205; Indels 58; Gaps 17;  
  
QY 13 GSKKQKQTEEDSDSPRS-----TAEALRAAPADNRV---TRVDPSCPSFRNPGTQVHED-Y 64  
Db 138 gakevlgkqdevevkkemvttatkgaavldghqldphkikvnyhv-----lqvgdely 190  
QY 65 DCTLQNTNIGNNNNFYIIQLLEE--GSRFFCWNRRGVGEVQSKMNHFTCLEL-AKKD 121  
Db 191 datlqntvngdnknkfyliqviesdaggsfmyvnrwgrvvgqdklhpsptrdqiye 250  
QY 122 FKXKFWKTKNKEERDRFVAQPNKYTLIEVQ-GEAESEAVVYKALSPQVDSGPVTVVK 180  
Db 251 fegkfhkntnnhwsdrknfkcyakkytwiemdyetek-----elekgsitdqik 300  
QY 181 PCSLDPATQNLITNIFSKEMFNAMTLMNLDVKMPLGKLTQOIARGFEALEEAMK 240  
Db 301 etkletriaqflislicnismmkqrmveigynaekiplgkirkatlikghvnlkrisdvis 360  
QY 241 NPTGDCQSLEELSSCFYTVIPHNF--RSRPPPIINSDDVLOAKKMDLLVLADIELAQTQL 298  
Db 361 k--adrrhleqitgefytviphdfgrkmrefiidtpqklkaklenvealgeieiatkll 418  
QY 299 AAPGEEEEKVEVPHPLDRDYQLRCQLLDGSESEYKAIQTYLKQT---GNSYRCPNL 355  
Db 419 -----edssdqddplyaryqklhcdftpleadsdeysmiksylnrthgkthsgytvdi 472  
QY 356 RHVWKVNRGEGRDQAHSKLGNRLWLHGTVNAVVAAILTSGLRIMPH---SGGRVCK 411  
Db 473 vqlfkvsrthgeterfqkfastrnrmllwhgsrlsnwagllsqglrriappeavptgymfgk 532  
QY 412 GIYFASSENSKASGYVTTMHCQGHQV---GYMFLGEVALGKEHHITIDDPSSLKSPPPGDFS 468  
Db 533 gvyfadmfsksany-----cyaseacrsyglilcevalgdmnellnadynnlpgkklr 587  
QY 469 VTARGOTEPDPAQDIELDGDFVVPVQGPVQCPFSKSSFSQSEYLIYKESQCRRLYL 528  
Db 588 skgvsggtapnm---veskvaddgvvvvplgepkqeps-krggillyneyivynvdqirmrv 643  
QY 529 LEIH 532  
Db 644 lhvn 647

RESULT 7  
AAR99642  
ID AAR99642 standard; Protein; 1013 AA.  
XX  
XX AAR99642;  
XX  
XX 31-OCT-1996 (first entry)  
XX  
XX Poly(ADP-ribose) polymerase contg. DNA-binding domain.  
XX  
XX PARP; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;  
KW tumour treatment; DNA repair; over-expression.  
XX  
XX Homo sapiens.  
XX  
XX WO9618737-A2.  
XX  
XX 20-JUN-1996.  
XX  
XX 15-DEC-1995; 95WO-DE01817.  
XX  
XX 16-DEC-1994; 94DE-4444949.  
XX  
XX (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
PA





```
XX 02-MAR-2000; 2000US-0517467.
XX (ISIS-) ISIS PHARM INC.
XX PI Popoff I, Cowsert LM;
XX XX
XX WPI; 2001-602570/68.
XX DR N-PSDB; AAS45582.
XX XX
XX Antisense compound useful for treating hyperproliferative,
XX PT neurological, inflammatory and autoimmune disorders and diabetes
XX PT inhibits human PARP -
XX XX
XX Example 13; Page 97-101; 168pp; English.
XX CC
XX CC The invention relates to antisense oligonucleotides targeted to human
XX CC PARP nucleic acid and inhibiting expression of human PARP. PARP
XX CC (Poly (ADP-ribose) polymerase plays an important role in chromatin
XX CC decondensation, DNA replication, DNA repair, gene expression, malignant
XX CC transformation, cellular differentiation and apoptosis. The antisense
XX CC oligonucleotide inhibitors are useful for inhibiting the expression of
XX CC PARP in human cells or tissues. They are also useful for treating a
XX CC human with a disease associated with PARP especially hyperproliferative
XX CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,
XX CC neurological (e.g. parkinsonism, meningitis-associated intracranial
XX CC complications and ischaemia), inflammatory and autoimmune disorders (e.g
XX CC arthritis) and diabetes. The present sequence is a PARP protein,
XX CC the cDNA encoding which was used to design the antisense
XX CC oligonucleotides.
XX XX
XX Sequence 1014 AA;

Query Match 24.9%; Score 700; DB 22; Length 1014;
Best Local Similarity 33.6%; Pred. No. 2.5e-53;
Matches 182; Conservative 102; Mismatches 219; Indels 38; Gaps 17;

QY 1 MAPKRKASVQTGSKKQROGTEEDSFRTAEALRAAPADNRVIRVDPCFSPFRNPQIQV 60
DB 493 vaprgkska---alskksqgvkeeginksekmkltlkggaavdpdsglehsahvlegk 549
QY 61 HEDYDCTLNQNTGNNNKFIYIQLLEEG--SRFFCWNRMGRVGEV-GOSKMNHFTCLEL 117
DB 550 gkvfsatlgldvkvgtstnsyvkqlleddkenrywifswgrvtvgsnkleqmpsked 609
QY 118 AKDFKKFKWEKTKNWEERDRVAOPNKYTLLEVOGEAESQAVVKALSPQVDSGPVRT 177
DB 610 aleqfmklyeektgnawhskn-ftkypkfkfyleid-ygqdeavvkl---tvpngtck 664
QY 178 VVRPCSLDPATONLIITNIFSKFMFNAMTLMNLDVKKMPLGKLTQKQIARGFEALEALEE 237
DB 665 lpkp-----vqdlkmifdvesmkkamveyeidlqkmpglgkskriqaaaysilsevqq 718
QY 238 AMKNPTGDSGLSEELSCFTVTIHPNFRSRPPPIINSDVLAQKMDMLVLADIELAOTL 297
DB 719 avsgssdsqll-dlsnrftllphdfgmkkppllnnadsqkvemldnllldieavysl 777
QY 298 QAAPGEEERKEVEVPPLDRDYOLLRCQLQLDSEGESEYKAIOYTLKOT-GNSYRCPNLR 356
DB 778 lrggsddsk-----qpIdvnyekltldikvdrdseaealfkyvknthattchsydlle 832
QY 357 --HWKVNREGEGRFOAHSKLGNRRLLHGTNTNVAVVAALITSLGRIMPH---SGGRVG 410
DB 833 vidifkiergeqcykpfqklnhrrllwhgstrtnfagilsglriapeaptygmfg 892
QY 411 KGIYFASENSKAGYVTTHMCGHGQVGYMFLGEVALGKEHHITIDPDSLKSPPPGDSVI 470
DB 893 kglyfadmvsksnanyhtsq--gdpiglillgevaiglmnyelk-hashisrlpkghskv 949
QY 471 ARQGTEDPAQDILELDDQPVVVRQGPVQPCPSFRKSSSPQSELYLYKESQCRLAYLLE 530
DB 950 glgkttddpsani--sldg--vdvpigtgis--sgvidtsllyneyivydiavnlkyllk 1004
```

```
QY 531 I 531
DB 1005 I 1005

RESULT 11
AAB66296
ID AAB66296 standard; Protein; 1014 AA.
XX
XX AAB66296;
XX
XX 05-APR-2001 (first entry)
XX
XX Human tankyrase2 related protein sequence SEQ ID NO: 137.
XX
XX Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;
XX KW inflammatory disorder.
XX
XX Homo sapiens.
XX
XX WO200100849-A1.
XX
XX 04-JAN-2001.
XX
XX 28-JUN-2000; 2000WO-US17827.
XX
XX 29-JUN-1999; 99US-0141582.
XX
XX (ICOS-) ICOS CORP.
XX
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;
XX
XX WPI; 2001-102896/11.
XX DR N-PSDB; AAF63954.
XX
XX New tankyrase2 polypeptides, useful for treating conditions mediated by
XX PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,
XX PT inflammatory and autoimmune disorders -
XX
XX Example 2; Page 207-210; 242pp; English.
XX
XX The present invention provides the protein and coding sequence for the
XX CC human tankyrase2 protein. This is found in two different versions,
XX CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has
XX CC polyADP-ribosylation activity and is involved in the modification of
XX CC TRF1, which is a telomere-specific binding protein. The regulation of
XX CC telomere length, in which TRF1 has a role, is linked to ageing and
XX CC cancer. The sequences are useful in the treatment of cancers and
XX CC inflammatory disorders.
XX
XX Sequence 1014 AA;

Query Match 24.9%; Score 700; DB 22; Length 1014;
Best Local Similarity 33.6%; Pred. No. 2.5e-53;
Matches 182; Conservative 102; Mismatches 219; Indels 38; Gaps 17;

QY 1 MAPKRKASVQTGSKKQROGTEEDSFRTAEALRAAPADNRVIRVDPCFSPFRNPQIQV 60
DB 493 vaprgkska---alskksqgvkeeginksekmkltlkggaavdpdsglehsahvlegk 549
QY 61 HEDYDCTLNQNTGNNNKFIYIQLLEEG--SRFFCWNRMGRVGEV-GOSKMNHFTCLEL 117
DB 550 gkvfsatlgldvkvgtstnsyvkqlleddkenrywifswgrvtvgsnkleqmpsked 609
QY 118 AKDFKKFKWEKTKNWEERDRVAOPNKYTLLEVOGEAESQAVVKALSPQVDSGPVRT 177
DB 610 aleqfmklyeektgnawhskn-ftkypkfkfyleid-ygqdeavvkl---tvpngtck 664
QY 178 VVRPCSLDPATONLIITNIFSKFMFNAMTLMNLDVKKMPLGKLTQKQIARGFEALEALEE 237
DB 665 lpkp-----vqdlkmifdvesmkkamveyeidlqkmpglgkskriqaaaysilsevqq 718
```







Query Match 24.8%; Score 698; DB 21; Length 570;  
 Best Local Similarity 33.9%; Pred. No. 1.5e-53;  
 Matches 185; Conservative 92; Mismatches 205; Indels 64; Gaps 18;

QY 13 GSKKQQTGEEDSFRSTAEALRA-APADNRVIRVDPSCFSPSRNPG---IQVHEDYDCTL 68  
 DB 54 gkankdrtdkqdesvkalllkgkap-----vdpctakvgkavhycegndvdyvml 106

QY 69 NOTNIGNNNKFFIOLLEGSR--FFCWNWRGVRGEVQSKNHFHC---LEDAKDKFK 123  
 DB 107 nqnlgnfnknyllqllledadgnfsvmwrgrvdkmgqshl--vacsgnlnkakeifq 164

QY 124 KKFWEKTKNWEERDRFVAQPNKYTLIEV-----QGEAESQEAIVKALSPQVDSGPV 175  
 DB 165 kfkldtknwnedrefekvpqkydmlqmdyatanqtqdeetkkeslksplkpe----- 218

QY 176 RTVVKPCSLDPATONLITNIFSKEMFNKAMTLMNLDVKMKPLGKLTQQTARFEALEAL 235  
 DB 219 -----sqldlrvgelklkicnvqameemmenkntkkgaplgltvagikagysgalkk 272

QY 236 EEAMKNPTGD-GQSELELSSCFVTIPIHFNFRSRPPPIINSPDLVQAKKMLVLADIELA 294  
 DB 273 edcir--agghgralmeacneftriphdfgrltpplrtqkelskigllealgieia 330

QY 295 QTLQAAPGEEEEKVEEVPHPDLDRDYQLLRQCLQLDLSGESEYKAIQIOTYKQT---GNSYR 351  
 DB 331 lkl-----vktelqpehpldqhyrnlcalrpldhesyefkvisqylqsthaphsdy 384

QY 352 CPNLRHWKVRREGDRFOAHKSLGNRLRHGHTNVAVVAAILTSGLRMPH----SGG 407  
 DB 385 tmtlldlfevekdekeaf--edlnhmlwhgmsnwgilshglrlappeapitgy 442

QY 408 RVCKGIYFASNSKAGYVYTHHCGHQVGMFLGEVAGLKEHHIITDPSLKSPPPGFD 467  
 DB 443 mfgkgyfyfamsksanycfasrl--kntglllsevalgqcnellleapkaegllqgkh 500

QY 468 SVIARGQTEPDPAQDIELDGQPVVPGPVQCPSPFKSSSF--SQSEYLIYKESQCRCL 525  
 DB 501 stkglgkmapssahfv--tlngs--cvplgpasdtgilnpdgytlnyneyivynpqvrm 556

QY 526 RYLLEI 531  
 DB 557 ryllkv 562

RESULT 15  
 ABL11480  
 ID AAB11480 standard; Protein; 570 AA.  
 AC AAB11480;  
 XX  
 DT 02-MAR-2001 (first entry)  
 XX  
 DE Human brain poly-ADP-ribose-polymerase protein.  
 XX  
 KW Human; poly ADP-ribose polymerase; PARP; neuroprotective; neurotropic; cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotrophic; anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment; antinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy; "2H-phthalazin-1-one derivative; neurodegenerative disease; tumor; neuronal damage; Alzheimer's disease; Huntington's disease; metastasis; Parkinson's disease; ischemic damage; microinfarction; sepsis; diabetes mellitus.  
 XX  
 OS Homo sapiens.  
 XX  
 PN DE19921567-A1.  
 XX  
 PD 16-NOV-2000.  
 XX  
 PF 11-MAY-1999; 99DE-1021567..  
 XX

PR 11-MAY-1999; 99DE-1021567.  
 XX (BADI ) BASF AG.  
 XX Lubisch W, Sadowski J, Kock M, Hoeger T;  
 PI WPI: 2001-032983/05.  
 DR N-PSDB; AAC82090.  
 XX  
 XX Drugs for inhibiting PARP or especially homologous enzymes comprising  
 PT 4-substituted phthalazinone derivatives, useful e.g. for treating  
 PT neurodegenerative disease, ischemic damage, tumors or diabetes -  
 XX  
 XX Example A; Page 12-13; 14pp; German.  
 XX  
 CC This invention describes novel 4-substituted 2H-phthalazin-1-one  
 CC derivatives (I) which are used for the treatment or prophylaxis of  
 CC diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;  
 CC EC 2.4.2.30) activity. The products of the invention have neurotropic,  
 CC neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,  
 CC cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,  
 CC immunosuppressive, antinflammatory, antirheumatic, antiarthritic,  
 CC antidiabetic. (I) are especially used for treating or preventing  
 CC neurodegenerative disease or neuronal damage (especially apoplexy or  
 CC with ischemia, trauma or massive bleeding, especially apoplexy or  
 CC spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or  
 CC Parkinson's disease), treating or preventing ischemic damage  
 CC (specifically renal damage after renal ischemia or during and after  
 CC kidney transplantation or heart damage after cardiac ischemia), treating  
 CC epilepsy, specifically generalized epileptic attacks (e.g. petit mal and  
 CC tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lobe  
 CC and complex partial attacks), treating microinfarction (e.g. during and  
 CC after heart valve replacement, aneurysm resectioning and heart  
 CC transplantation), revascularization of critically constricted coronary  
 CC arteries (e.g. after PCRA or by-pass operations) or peripheral arteries  
 CC (e.g. leg arteries), treating acute myocardial ischemia and damage during  
 CC or after its mechanical or drug-induced lysis and treating tumors and  
 CC their metastasis, sepsis and septic shock, inflammatory and rheumatic  
 CC disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)  
 CC inhibit PARP (i.e. the known form designated PARP1), they especially  
 CC selectively and strongly inhibit PARP homologs, specifically the homolog  
 CC PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.  
 CC with K<sub>i</sub> values of 1-20 nM) and high selectivity for PARP2 relative to  
 CC PARP1 (generally by a factor of more than 5).  
 XX  
 XX Sequence 570 AA;

Query Match 24.8%; Score 698; DB 22; Length 570;  
 Best Local Similarity 33.9%; Pred. No. 1.5e-53;  
 Matches 185; Conservative 92; Mismatches 205; Indels 64; Gaps 18;

QY 13 GSKKQQTGEEDSFRSTAEALRA-APADNRVIRVDPSCFSPSRNPG---IQVHEDYDCTL 68  
 DB 54 gkankdrtdkqdesvkalllkgkap-----vdpctakvgkavhycegndvdyvml 106

QY 69 NOTNIGNNNKFFIOLLEGSR--FFCWNWRGVRGEVQSKNHFHC---LEDAKDKFK 123  
 DB 107 nqnlgnfnknyllqllledadgnfsvmwrgrvdkmgqshl--vacsgnlnkakeifq 164

QY 124 KKFWEKTKNWEERDRFVAQPNKYTLIEV-----QGEAESQEAIVKALSPQVDSGPV 175  
 DB 165 kfkldtknwnedrefekvpqkydmlqmdyatanqtqdeetkkeslksplkpe----- 218

QY 176 RTVVKPCSLDPATONLITNIFSKEMFNKAMTLMNLDVKMKPLGKLTQQTARFEALEAL 235  
 DB 219 -----sqldlrvgelklkicnvqameemmenkntkkgaplgltvagikagysgalkk 272

QY 236 EEAMKNPTGD-GQSELELSSCFVTIPIHFNFRSRPPPIINSPDLVQAKKMLVLADIELA 294  
 DB 273 edcir--agghgralmeacneftriphdfgrltpplrtqkelskigllealgieia 330

QY 295 QTLQAAPGEEEEKVEEVPHPDLDRDYQLLRQCLQLDLSGESEYKAIQIOTYKQT---GNSYR 351

```

Db 331 ikl-----vkteiqspehpldqnyrnlhcalrpldhesyefkvisqyiqsthapthsdY 384
QY 352 CPNLRHWKVNREGEDRFOAHSKLGNRRLWHGTNNVAVVAAILTSGLRIMPH-----SGG 407
Db 385 tmtlldlfevekdgeakefr--edlhnrmllwhgmsnwnvgilshgllriapeapitgy 442
QY 408 RVKGGIYFASSENSKSAGYVTTMHCGGHQVGYMFLGEVALGKEHHITIDDPKSPPPGFD 467
Db 443 mfgkgiyfadmsksksanycfasrl--kntgllllsevalgqcnelleanpkaegllqgkh 500
QY 468 SVIARGQTEPDPAQDIELELDGQPVVPOGPPVQCPSFKSSSF--SQSEYLIYKESQCRL 525
Db 501 stkglgkmapssahfv--tngs--tvpigpasdtgilnpdgytlnyneyivynpnqvm 556
QY 526 RYLLEI 531
Db 557 ryllkv 562

```

Search completed: August 29, 2002, 07:57:19  
Job time: 296 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:59:33 ; Search time 69.02 Seconds  
(without alignments)  
742.040 Million cell updates/sec

Title: US-09-701-586b-8  
Perfect score: 2813  
Sequence: 1 MAPKKKASVQTEGSKRQGG.....EYLIYKESQCLRLYLEIHL 533  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1975	70.2	459	2 T08713	NAD+ ADP-ribosyltr
2	717	25.5	653	2 T03656	probable NAD+ ADP-
3	705	25.1	1014	1 A29725	NAD+ ADP-ribosyltr
4	702	25.0	1016	1 JS0428	NAD+ ADP-ribosyltr
5	686.5	24.4	998	2 S31735	NAD+ ADP-ribosyltr
6	686.5	24.4	1011	1 JH0581	NAD+ ADP-ribosyltr
7	683.5	24.3	635	2 T01311	NAD+ ADP-ribosyltr
8	683.5	24.3	1013	1 S04200	NAD+ ADP-ribosyltr
9	682.5	24.3	500	2 S26057	NAD+ ADP-ribosyltr
10	645.5	22.9	996	1 S42208	NAD+ ADP-ribosyltr
11	621.5	22.1	983	2 T51353	NAD+ ADP-ribosyltr
12	621.5	22.1	1009	2 C84719	probable poly (ADP
13	619	22.0	994	1 A47474	NAD+ ADP-ribosyltr
14	584.5	20.8	969	2 T03657	NAD+ ADP-ribosyltr
15	475.5	16.9	727	2 T18600	hypothetical prote
16	396.5	14.1	538	2 T20414	hypothetical prote
17	193	6.9	135	2 PN0494	NAD+ ADP-ribosyltr
18	180	6.4	2004	2 D88948	protein zK1005.1 (
19	125.5	4.5	5175	2 T20992	hypothetical prote
20	125.5	4.5	5198	2 T43290	hemichitin precurs
21	120.5	4.3	181	2 T03058	NAD+ ADP-ribosyltr
22	119.5	4.2	954	2 E86174	protein F19p19.26
23	118.5	4.2	1092	2 T33717	carbamoyl-phosphat
24	114.5	4.1	1203	2 D86625	subunit A of ATP-d
25	112.5	4.0	935	2 E96806	hypothetical prote
26	111.5	4.0	359	2 T12340	hypothetical prote
27	110	3.9	1223	2 T17345	hypothetical prote
28	108	3.8	949	2 D97781	hypothetical prote
29	106.5	3.8	638	2 C83503	exotoxin A precurs

30	106.5	3.8	1031	1 A38713	kinesin heavy chai
31	106.5	3.8	1435	2 S96632	regulatory protein
32	106	3.8	769	2 E96613	hypothetical prote
33	105.5	3.8	4162	2 T42633	connectin/titin -
34	105	3.7	511	2 S38905	hypothetical prote
35	105	3.7	3187	2 JC5837	364K Golgi complex
36	104.5	3.7	963	1 A41919	kinesin heavy chai
37	104.5	3.7	1173	2 T25539	hypothetical prote
38	104	3.7	2484	2 T26216	hypothetical prote
39	104	3.7	2607	2 T26215	hypothetical prote
40	103.5	3.7	638	2 A30347	exotoxin A precurs
41	102.5	3.6	1280	2 T00365	hypothetical prote
42	102	3.6	881	2 T84737	kinesin heavy chai
43	101.5	3.6	1244	2 A05218	hypothetical prote
44	101	3.6	783	2 T18421	hypothetical prote
45	101	3.6	3075	2 S14458	laminin alpha-1 ch

ALIGNMENTS

RESULT 1

T08713  
NAD+ ADP-ribosyltransferase homolog DKFp566G0224.1 - human (fragment)  
C:Species: Homo sapiens (man)  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Nov-2000  
C:Accession: T08713  
R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
submitted to the Protein Sequence Database, May 1999  
A:Reference number: Z16472  
A:Accession: T08713  
A:Molecule type: mRNA  
A:Residues: 1-459 <ANS>  
A:Cross-references: EMBL:AL050034  
A:Experimental source: fetal kidney; clone DKFp566G0224  
C:Genetics:  
A:Note: DKFp566G0224.1  
C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 70.2%; Score 1975; DB 2; Length 459;  
Best Local Similarity 81.5%; Pred. No. 2.3e-135;  
Matches 379; Conservative 32; Mismatches 44; Indels 10; Gaps 4;

Qy	73	IGNNNFYIIQLLEGRSFF-CHNRWCRVCEVGSQKMNHFTCLDEAKDKKKFKWEKTK	131
Db	1	IENNNFYIIQLLDNSNRFFTCWNRWGRVGEVGSQKINHFTRLEDAKKDKFEKPREKTK	60
Qy	132	NKWEERDFVAOPNKYTLIEVQGEAESQEAQVAVKALSPQVDSGPVRTV---VKPCLSDPAT	188
Db	61	NNWAERDFVSHPKYTLIEVQAEAEQAVVK-----VDRGPVRTVTKRVQPSLDPAT	115
Qy	189	QNLITNIFSKEMFNAMTLMNLDVKKMPGLKLTQQTARGFEALAEAEAMKNFTDGGQS	248
Db	116	QKLITNIFSKEMFNAMTLMNLDVKKMPGLKLSQQTARGFEALAEAEALGPTDGGQS	175
Qy	249	LEELSSCFYTVIPNFGSRPPPPINSDVLOAKKMDLLVLADIQAOTLQAPGEEKVK	308
Db	176	LEELSSHFYTVIPNFGSQPPPIINSPQLLQAKKMDLLVLADIQAALQALQV-SEQEKTV	234
Qy	309	EEVPHPLDRDYQLLRCLQLDLSGESEYKATQTYLKOTGNSYRCPNLRHVWKNVREGSD	368
Db	235	EEVPHPLDRDYQLLRCLQLDLSGAPEYKVQTYLEQTGSNHRCPFTQHIWKNVQEGED	294
Qy	369	RFQAHSLGNLRRLLWGHGTVAVVAAILTSLGRIMPHSGRVRGKGIYFASENSKSAGVYTT	428
Db	295	RFQAHSLGNLRRLLWGHGTVAVVAAILTSLGRIMPHSGRVRGKGIYFASENSKSAGVYIG	354
Qy	429	MKCGGHGVGMFLGCEVAGLKEHHITIDDPKLSKPPGFDSDVIARGQTEPDPAQDIEELD	488
Db	355	MKCGAHGVGMFLGCEVAGLKEHHITIDNPDKLSKPPGFDSDVIARGHTEPDPTQDTEELD	414
Qy	489	GQPVVPGPVQCPQSFSSFSQSEYLIYKESQCLRLYLEIHL	533



A:Molecule type: mRNA  
A:Residues: 381-420:682-710 <SCH>  
R:Yokoyama, Y.; Kawamoto, T.; Mitsunuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Terashima, J. Biochem. 194, 521-526, 1990  
A:Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.  
A:Reference number: S14010; MUID:91099327  
A:Accession: S14010  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-95 <YOK>  
A:Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA39606.1; PID:g825702  
C:Comment: This protein can ADP-ribosylate itself as well as other proteins.  
C:Genetics:  
A:Gene: GDB:ADPRT; PPOL  
A:Cross-references: GDB:119508; OMIM:173870  
A:Map position: 1q41-1q42  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match 25.1%; Score 705; DB 1; Length 1014;  
Best Local Similarity 33.8%; Pred. No. 8.1e-43; Indels 40; Gaps 17;  
Matches 183; Conservative 101; Mismatches 218; Mismatches 218; Indels 40; Gaps 17;

QY 1 MAPKKASVOTEGSKKORQTEEDSPRSTAEALRAAPADNRVIRVDPSPCFSPSRNPQIV 60  
DB 493 VAPRGKSGA---ALSRSKQGVKEEGINKSEKMKLTLKGAAVDPDPSGLEHSAHVLEKG 549  
QY 61 HEDYDCTLTNTGNNNKYYIQLLEB--SRFFCWNRRWGRVGEV-GQSKMNHFTCLE 117  
DB 550 GKVPFATGLVDIVKGTNSYKQLLEDDEKRYWIFRSWGRVGTIGSNKLEQMSKED 609  
QY 118 AKDKPKKFKWEKTKNKEERDRFVAOPNKYTLLEVOGEAESEAVVKALSPQVDSGPVRT 177  
DB 610 AIEHFMKLYEEKTCGNNAHNSKN-FTKYPKKFYPLEID-YGQDEEAVKKL---TVNPGTKSK 664  
QY 178 VVKPCSLDPATQNLITNIFSKEMFNAMTLNLDVKKMPLGLKLTQKQIARGFALEALEE 237  
DB 665 LPKP-----VQDLIKMIFDVESMKKAMVEYEDLQKMPGLKLSKRQIQAAYSITLSEVQQ 718  
QY 238 AMKNPTGQGSLELSCTFYVPHNFGSRPPPIPSDVLQAKKMLLVADIQLAQTLL 297  
DB 719 AVSGSSDSQIL-DLSNRFYTLPHDFGKMKPPLNNAVSQAKVEMLDNLDDIEAVSL 777  
QY 298 QAAPGEEEEKVEEVPHPDLRDYQLLRQQLDLSGESEYKAIQTYLQKQ---TCNSYRCP 353  
DB 778 LRGSDDSSK-----DPIDVNYEKLKTDIKVVDKSEAEIIRKYVKNTHATTNATYDL- 831  
QY 354 NLRHVKNRREGEDRFOAHSKLGKLNRLWHTGNVAVVAAILTSGLRIMPH-----SGGRV 409  
DB 832 EVIDIFKIEREGECQRYKPKQLHNRLLWHSRTTNFAGILSQGLRIAPPEAPVTGYMF 891  
QY 410 KGIFYASENSKSNAGYVTHMCGHQGVYMFLEVALGKEHHITIDDPSPKSPPPGDSV 469  
DB 892 KGIFYADWYKSNANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGRHSV 948  
QY 470 IARGOTEPDPAQDIELELDQPVVPPQVPCPSKSSFSOSEYLIYKESQCLRYLL 529  
DB 949 KGIGKTTTPDSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEYIYVDIAQVNLKYL 1003  
QY 530 EI 531  
DB 1004 KL 1005

RESULT 4  
JS0428  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine  
C:Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
C:Accession: JS0428; S00328; A30458  
R:Saito, I.

submitted to JIPID, February 1990  
A:Reference number: JS0428  
A:Accession: JS0428  
A:Molecule type: mRNA  
A:Residues: 1-1016 <SAI>  
A:Experimental source: thymus  
R:Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; T  
Eur. J. Biochem. 171, 571-575, 1988  
A:Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int  
A:Reference number: S00328; MUID:88151954  
A:Accession: S00328  
A:Molecule type: mRNA  
A:Residues: 648-714; 838-904 <TAN>  
A:Cross-references: EMBL:X06986  
A:Accession: A30458  
A:Molecule type: protein  
A:Residues: 658-685; 689-696; 893-901 <TA2>  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-100  
F:21-51/Region: zinc finger  
F:128-165/Region: zinc finger  
F:200-220/Region: helix-turn-helix motif  
F:224-231/Region: nuclear location signal  
F:250-270/Region: helix-turn-helix motif  
F:494-501/Region: nucleotide-binding motif A (P-loop)  
F:890-903/Region: nucleotide binding #status predicted

Query Match 25.0%; Score 702; DB 1; Length 1016;  
Best Local Similarity 34.2%; Pred. No. 1.3e-42;  
Matches 187; Conservative 95; Mismatches 215; Indels 50; Gaps 20;

QY 1 MAPKKASVOTEGSKKORQTEEDSPRSTAE---ALRAAPADNRVIRVDPSPCFSPSRNP 57  
DB 495 VGRGKSGAAP--SKRSKGPVKEGTNKRKMKLTUKGAA-----VDPDGLSEHNAH 546  
QY 58 I--QVHEDYDCTLTNTGNNNKYYIQLLEB--SRFFCWNRRWGRVGEV-GQSKMNH 112  
DB 547 VLEKGGKVFSAATGLVDIVKGTNSYKQLLEDDEKRYWIFRSWGRVGTIGSNKLEQ 606  
QY 113 TCLEDAKDKPKFKWEKTKNKEERDRFVAOPNKYTLLEVOGEAESEAVVKALSPQVDS 172  
DB 607 PSKEDATEHEFMKLYEEKTCGNNAHNSKN-FTKHPKKFYPLEID-YGQDEEAVKKL---TVNP 661  
QY 173 GPVTVVVKPCSLDPATQNLITNIFSKEMFNAMTLNLDVKKMPLGLKLTQKQIARGFEAL 232  
DB 662 GTSKSLPKP-----VQNLKIMIFDVESMKKAMVEYEDLQKMPGLKLSKRQIQAAYSIL 715  
QY 233 EALFEAMKNPTGQGSLELSCTFYVPHNFGSRPPPIPSDVLQAKKMLLVADI 292  
DB 716 SEVQALSQGSSDSHIL-DLSNRFYTLPHDFGKMKPPLNNAVSQAKVEMLDNLDDIE 774  
QY 293 LAOTLQAAPGEEEEKVEEVPHPDLRDYQLLRQQLDLSGESEYKAIQTYLQKQ---TGN 348  
DB 775 VAYSLLRGSSDDSSK-----DPIDVNYEKLKTDIKVVDKSEAEIIRKYVKNTHATTN 829  
QY 349 SYRCPNLRHVKNRREGEDRFOAHSKLGKLNRLWHTGNVAVVAAILTSGLRIMPH--- 404  
DB 830 AYDL-EVVDIFKIEREGESQRYKPKQLHNRLLWHSRTTNFAGILSQGLRIAPPEAPV 888  
QY 405 SGRVKGIFYASENSKSNAGYVTHMCGHQGVYMFLEVALGKEHHITIDDPSPKSPPP 464  
DB 889 TGYMFGKIFYADWYKSNANYCHTSQ--GDPIGLILLGEVALGNMYELK-HARHISKLPK 945  
QY 465 GFDSVIARGOTEPDPAQDIELELDQPVVPPQVPCPSKSSFSOSEYLIYKESQCR 524  
DB 946 GKHSVKGLGKTTTPDSASI--TVDDG--VEVPLGTGIS-SGVNDTCLLYNEYIYVDIAQVH 1000  
QY 525 LRYLLEI 531  
DB 1001 LKYLKLL 1007





**RESULT** 8



Db 472 KKQKLPFDKFKYKIEDTSESIVTVKVKGRSAVHEAS-----GLQERCHILEDGNSYNT 524  
QY 67 TLNQTNGNNNNFYIIQLLEE--GSRFFCWNWRGVRG--EVGSKMNHFTCLDEAKKDF 122  
Db 525 TLSMSDLSTGINSYIIQLIQEDKSDCYVFRKMGVRGNEKIGKNKVEEMS--KSDAVHEF 583  
QY 123 KKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQAVVYKALSPOVDSGPVRTV 179  
Db 584 KRLEKTGNTWESWEGKTNFQKPGKFLPLDIDYGVNKQVA-----KKEPFTQ-- 632  
QY 180 KPCSLDPATONLTINIFSKEMFNAMTLMNLDVKMPLGLTKLTKQOIARGFEAL----EAL 235  
Db 633 -SSNLAPSLIELKMLFDVETYSAMMEFEINMSEMPGLKSKHNIQKGFALTEIQRLL 691  
QY 236 EEMKNPTGQGSLEELSSCFYIVIPNFGRSRPPPIINSDVLAQKMDMLLVADIELAQ 295  
Db 692 TESDPQPTKESLVDASNRFFTMIP-----SIHPHIIRDEDDFKSKVLMLEALQDIEIAS 747  
QY 296 TLOAAPGEEBEKVEEVPHPDLDRDYQLLRCOLQLDSESEYKATQTYLKOTGNSYRCP-- 353  
Db 748 RI-----VGFVDVSTES-----LDDKYKLLHCDISPLPHDSEYRLIEKYL-----NTHAPTH 796  
QY 354 -----NLRHWKVNREGEGRFQAH--SKLGNRRLLHGTNVAVVAAILTSGLRI----MP 403  
Db 797 TENSLEEEVFALEREGEFDKYAPHREKLGKMLLHGSRLTNFVGLINQGLRIAPPEAP 856  
QY 404 HSGRVRGKGIYFASSENSKSGAGYVTTMCHGGHOVGYMFLGEVALGKEHHITDDPSLSKPP 463  
Db 857 ATGYMFGKGIYFADLVSKSAQCYT--CKKNPVGLMILLSEVALGEIHELT--KAKYMDKPP 913  
QY 464 PGDSVITARGOTEPDPAQDIELELDGQPVVPOGPPVQCPSPKSSFSQSEYLIYKESQC 523  
Db 914 RGRHSTKGLGKKVP---QDSEFAKWRGDVTPCGKPV--SKVRASELMYNEYIYDIAQV 969  
QY 524 RLRYLLEI 531  
Db 970 KLOFLKV 977

RESULT 12  
C84719  
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C84719  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: C84719  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1009 <STO>  
A:CROSS-references: GB:AE002093; NID:g4432827; PIDN:AAD20677.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g31320  
A:Map position: 2  
C:Superfamily: NAD+ ADP-riboseyltransferase

Query Match 22.1%; Score 621.5; DB 2; Length 1009;  
Best Local Similarity 32.7%; Pred. No. 9.1e-37;  
Matches 179; Conservative 87; Mismatches 209; Indels 73; Gaps 21;

QY 15 KQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPSRNPQIYV-----ED----YDC 66  
Db 498 KKQKLPFDKFKYKIEDTSESIVTVKVKGRSAVHEAS-----GLQERCHILEDGNSYNT 550  
QY 67 TLNQTNGNNNNFYIIQLLEE--GSRFFCWNWRGVRG--EVGSKMNHFTCLDEAKKDF 122

Db 551 TLSMSDLSTGINSYIIQLIQEDKSDCYVFRKMGVRGNEKIGKNKVEEMS--KSDAVHEF 609  
QY 123 KKFWEKTKN---KWEERDRFVAQPNKYTLIEVQGEAESQAVVYKALSPOVDSGPVRTV 179  
Db 610 KRLEKTGNTWESWEGKTNFQKPGKFLPLDIDYGVNKQVA-----KKEPFTQ-- 658  
QY 180 KPCSLDPATONLTINIFSKEMFNAMTLMNLDVKMPLGLTKLTKQOIARGFEAL----EAL 235  
Db 659 -SSNLAPSLIELKMLFDVETYSAMMEFEINMSEMPGLKSKHNIQKGFALTEIQRLL 717  
QY 236 EEMKNPTGQGSLEELSSCFYIVIPNFGRSRPPPIINSDVLAQKMDMLLVADIELAQ 295  
Db 718 TESDPQPTKESLVDASNRFFTMIP-----SIHPHIIRDEDDFKSKVLMLEALQDIEIAS 773  
QY 296 TLOAAPGEEBEKVEEVPHPDLDRDYQLLRCOLQLDSESEYKATQTYLKOTGNSYRCP-- 353  
Db 774 RI-----VGFVDVSTES-----LDDKYKLLHCDISPLPHDSEYRLIEKYL-----NTHAPTH 822  
QY 354 -----NLRHWKVNREGEGRFQAH--SKLGNRRLLHGTNVAVVAAILTSGLRI----MP 403  
Db 823 TENSLEEEVFALEREGEFDKYAPHREKLGKMLLHGSRLTNFVGLINQGLRIAPPEAP 882  
QY 404 HSGRVRGKGIYFASSENSKSGAGYVTTMCHGGHOVGYMFLGEVALGKEHHITDDPSLSKPP 463  
Db 883 ATGYMFGKGIYFADLVSKSAQCYT--CKKNPVGLMILLSEVALGEIHELT--KAKYMDKPP 939  
QY 464 PGDSVITARGOTEPDPAQDIELELDGQPVVPOGPPVQCPSPKSSFSQSEYLIYKESQC 523  
Db 940 RGRHSTKGLGKKVP---QDSEFAKWRGDVTPCGKPV--SKVRASELMYNEYIYDIAQV 995  
QY 524 RLRYLLEI 531  
Db 996 KLOFLKV 1003

RESULT 13  
A47474  
NAD+ ADP-riboseyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: A47474  
R:Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993  
A:Title: Cloning of cDNA encoding Drosophila poly (ADP-ribose) polymerase: leucine zip  
A:Reference number: A47474; MUID:93234521  
A:Accession: A47474  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-994 <UCH>  
A:CROSS-references: GB:D13806; GB:D13807; GB:D13808; NID:g303545; PIDN:BAA02964.1; PI  
A:Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIP:129704)  
C:Genetics:  
A:Gene: FlyBase:Parp  
A:CROSS-references: FlyBase:FBgn0010247  
C:Superfamily: NAD+ ADP-riboseyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match 22.0%; Score 619; DB 1; Length 994;  
Best Local Similarity 31.5%; Pred. No. 1.4e-36;  
Matches 168; Conservative 100; Mismatches 222; Indels 44; Gaps 17;

QY 16 KQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPSRNPQIYV-----VHED----YDCT 67  
Db 481 KSRIPKETTKSLNSYITKSMVPV-SRTFKVKDGLAVDPDPSGLEDIAHVYVDSNNKYSW 539  
QY 68 LNQTINNNNNKYIQLL--EESRFFCWNWRGVRG--EVGSKMNHFTCLDEAKKDFK 124  
Db 540 LGTFDIQRNNNSYIKVQLLKADKKKIYIFKSWGRTIGNIGNSKLEEDFTSESARNFKE 599  
QY 125 KFWKTKNKNKEERDRFVAQPNKYTLIEVQGEAESQAVVYKALSPOVDSGPVRTVVKCSL 184  
Db 600 IYADKTGNEYEQRDNFVKRTGRMYPIEQ--YDDQKLVKHSHFFTS-----KL 647



Search completed: August 29, 2002, 07:59:37  
Job time: 284 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:59:37 ; Search time 69.02 Seconds  
(without alignments)  
735.079 Million cell updates/sec

Title: US-09-701-586B-10  
Perfect score: 2789  
Sequence: 1 MAPKKASVQTGSKKQROG.....EYLIYKESQCLRLYLEIHL 528

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71.\*

1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1987.5	71.3	459	2 T08713	NAD+ ADP-ribosyltr
2	719.5	25.8	653	2 T03656	probable NAD+ ADP-
3	707.5	25.4	1014	1 A29725	NAD+ ADP-ribosyltr
4	704.5	25.3	1016	1 JS0428	NAD+ ADP-ribosyltr
5	691.5	24.8	635	2 T01311	NAD+ ADP-ribosyltr
6	689.5	24.7	1011	1 JH0581	NAD+ ADP-ribosyltr
7	688.5	24.7	1013	1 S04200	NAD+ ADP-ribosyltr
8	686.5	24.6	500	2 S26057	NAD+ ADP-ribosyltr
9	681.5	24.4	998	2 S31735	NAD+ ADP-ribosyltr
10	652.5	23.4	996	1 S42208	NAD+ ADP-ribosyltr
11	624.5	22.4	983	2 T51353	NAD+ ADP-ribosyltr
12	624.5	22.4	1009	2 C84719	probable poly (ADP
13	621.5	22.3	994	1 A47474	NAD+ ADP-ribosyltr
14	584.5	20.9	969	2 T03657	NAD+ ADP-ribosyltr
15	481.5	17.2	727	2 T18600	hypothetical prote
16	389.5	13.9	538	2 T20414	hypothetical prote
17	193.5	6.9	135	2 P80494	NAD+ ADP-ribosyltr
18	172.5	6.2	2004	2 D88948	protein ZK1005.1
19	121.5	4.4	1223	2 T17345	hypothetical prote
20	121.5	4.3	1092	2 T33717	carbamoyl-phosphat
21	120.5	4.3	181	2 T30358	NAD+ ADP-ribosyltr
22	120.5	4.3	5175	2 T20952	hypothetical prote
23	120.5	4.3	5198	2 T43290	hemocentin precurs
24	118.5	4.2	935	2 E96806	hypothetical prote
25	117.5	4.2	954	2 E86174	protein F19P19.26
26	112.5	4.0	1203	2 D86625	subunit A of Atp-d
27	111.5	4.0	359	2 T12540	hypothetical prote
28	107.5	3.9	638	2 C83503	exotoxin A precurs
29	106.5	3.8	646	2 T38022	probable GTP-bindi

30	105.5	3.8	3075	2 SI4458	laminin alpha-1 ch
31	104.5	3.7	2285	2 T12796	probable transglyc
32	103.5	3.7	638	2 A30347	exotoxin A precurs
33	103.5	3.7	357	1 Q0MSLL	retrovirus-related
34	103.5	3.7	1031	1 A38713	kinesin heavy chai
35	103.5	3.7	1435	2 S69632	regulatory protein
36	102.5	3.7	354	2 D90187	oxidoreductase [im
37	102.5	3.7	612	1 WDBPT5	gene D13 protein -
38	102.5	3.7	2484	2 T26216	hypothetical prote
39	102.5	3.7	2607	2 T26215	hypothetical prote
40	101.5	3.6	295	2 F83412	probable transcrip
41	101.5	3.6	398	2 E71424	hypothetical prote
42	101.5	3.6	424	2 H85175	hypothetical prote
43	101.5	3.6	436	2 C90043	IgG-binding protei
44	101.5	3.6	769	2 E96613	hypothetical prote
45	101.5	3.6	1108	2 E71104	probable cell divi

ALIGNMENTS

RESULT 1

T08713

NAD+ ADP-ribosyltransferase homolog DKF2p566G0224.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Nov-2000

C:Accession: T08713

R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16472

A:Accession: T08713

A:Molecule type: mRNA

A:Residues: 1-459 <ANS>

A:Cross-references: EMBL:AL050034

A:Experimental source: fetal kidney; clone DKF2p566G0224

C:Genetics:

A:Note: DKF2p566G0224.1

C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 71.3%; Score 1987.5; DB 2; Length 459;  
Best Local Similarity 82.4%; Pred. No. 9e-136;  
Matches 379; Conservative 32; Mismatches 44; Indels 5; Gaps 3;

Qy	73	IGNNNKYYIIQLLEEGSRFF-CHNRWGRVGEVGSKNHHFTCLDEDAKDKFKKFEWTKT	131
Db	1	IENNNKYYIIQLLQDSNRFFTCNWRWGRVGEVGSKNHHFTCLDEDAKDKFEKFRKTK	60
Qy	132	NKWEERDRFVAOPNKYTLIEVQGEAEQAVVVKVDSGPVRTV---VKPCSLDPATONLIT	188
Db	61	NNWAERDFVSHPGKYTLIEVQAEAEQAVVVKVDRGPVRTVTKRVQPCSLDPATOKLIT	120
Qy	189	NFSKEMFKNAWTLMLNDVKKMPLGKLTQQIARGFEALEALEEAMKNPTDGOQSLEELS	248
Db	121	NFSKEMFKNTMALMDLVKKMPLGKLSKQIARGFEALEALEEALKGPTDGOQSLEELS	180
Qy	249	SCFTYVIPHNGRSRPPPIPSDVLQAKMLLVLADELIAQTLQAAPEEEKEVEEYVPH	308
Db	181	SHEFTVIPHNGFSQPPPIPSPELLQAKMLLVLADELIAQALQAV-SEQEKTVEEYVPH	239
Qy	309	PLDRDYQLLRQLQDLSGESEYKAIQYLTQSGNSYRCPNLRHVKNVKNREGEGRFOAH	368
Db	240	PLDRDYQLKQLQDLSGAPEYKVIQYLTQSGNHRCPTQLHVLKVNQEGEEDRFOAH	299
Qy	369	SKLGNRRLLHGTNNVAVVAAILTSGLRIMPHSGGVRGVIYFASSENSKSGAVYVTTHHCGG	428
Db	300	SKLGNRRLLHGTNNVAVVAAILTSGLRIMPHSGGVRGVIYFASSENSKSGAVYVIMKCGA	359
Qy	429	HOVGYMFLGEVALGKEHHITDDPSLKSPPGFDSVIARGQTEPPDPAQDIELELQGPVW	488
Db	360	HVGYMFLGEVALGREHHINTDNPSPKSPPGFDSVIARGHTPEPDPTQDTELELQGPVW	419
Qy	489	VPQGPVQCFKSSFSQSEYLIYKESQCLRLYLEIHL	528





A:Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.  
A:Reference number: S14010; MUID:91099327  
A:Accession: S14010  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-95 <YOK>  
A:Cross-references: EMBL:X56140; NID:g35286; PIDN:CAA9606.1; PID:g825702  
C:Comment: This protein can ADP-ribosylate itself as well as other proteins.  
C:Genetics:  
A:Gene: GDB:ADPRT; PPOL  
A:Cross-references: GDB:l19508; OMIM:173870  
A:Map position: lq41-lq42  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; DNA repair; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match 25.4%; Score 707.5; DB 1; Length 1014;  
Best Local Similarity 34.0%; Pred. No. 7.6e-43;  
Matches 183; Conservative 101; Mismatches 216; Indels 39; Gaps 17;

Qy	1	MAPKKASVQTEGSKKQKQGTEDSFRSTAEALRAAPADNRVIRVDPSPCFPSRNPQIQV	60
Db	493	VAPRKSGA---ALSKKSGQVKKEGINKSEKRMKLTILKGAADVDPDSGLHSAHLEKG	549
Qy	61	HEDYDCTLNQTNIGNNNKFFIYIQLLEEG--SRFFCMNRWRGVRGEV--GQSKMNHFTCL	117
Db	550	GKVFSAITGLVDIVKGTNSYKQLQLEDDKENRYWIFRSWGRVGVIGSNKLEQMPKSD	609
Qy	118	AKKDFKKFWKTKNKEWRDRFVAQPNKYIILIEVOGAESEAVVK--VDSGPVRTVVK	175
Db	610	ATIEHFMKLYEKTGNAMHSKN-FTKYPKKFYPLEID-YGODEEAVVKLTVPNGTKSKLPK	667
Qy	176	PCSLDPATQNLITNIFSKEMFNKNAWTLNLDVKKMPLGLKLTQKQIARGFEALEEAMK	235
Db	668	P-----VQDLIKMIFDVESMKKAMVEYIDLQKMPGLKSKRQIQAAYSILSEVQQA	721
Qy	236	NPTDGGQSLSELSFCYTVIPHNFGRSPPPIPSDVLOAKKMDLLVLADIQLAA	295
Db	722	QGSDDSQIL-DLSNRFYTLIPHDFGMKPPILLNADSVQAKVEMLDLDDIEVAYSLLRG	780
Qy	296	PGEEEKVEVPHPLDRIYQLLRQLQLDLSGSEYKAIQIYTLQKQ----TGNSTRCPNLR	351
Db	781	GSDSSK-----DPIDVNYEKLKTDIKVYDRDSEAEIIRKYVKNTHATTNAYDL-EVI	834
Qy	352	HWKVNREGRDFOAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGRVCVK	407
Db	835	DIKIEREGECQRYKPFQKLNRRLLWHGSRRTTNFAGILSQGLRIAPPEAPVTGMFGK	894
Qy	408	IYFASENSKSAGYVYTMHCGHGVYMFGLGEVALGKEHHITIDPSPKSPPPGDSVIAR	467
Db	895	IYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKGRHSVKGL	951
Qy	468	GQTEPDPDAQDIELELDGQPVVPPQVPPVQCFPSKSSFSQSEYLIYKESQCLRYLLEI	526
Db	952	GKTPDPDSANI--SLDG--VDVPLGTGIS-SCVNDTSLLYNEYIVYDIAQVNLKYLKL	1005

RESULT 4  
JS0428  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - bovine  
N:Alternate names: ADP-ribosyltransferase (polymerizing); poly(adenosine diphosphate rib  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 19-Jan-2001  
C:Accession: JS0428; S00328; A30458  
R:Saito, I.  
submitted to JIPID, February 1990  
A:Reference number: JS0428  
A:Accession: JS0428  
A:Molecule type: mRNA  
A:Residues: 1-1016 <SAI>  
A:Experimental source: Thymus  
R:Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, N.; Tanaka, H.; Take  
Eur. J. Biochem. 171, 571-575, 1988

A:Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int  
A:Reference number: S00328; MUID:88151954  
A:Accession: S00328  
A:Molecule type: mRNA  
A:Residues: 648-714; 838-904 <TAN>  
A:Cross-references: EMBL:X06986  
A:Accession: A30458  
A:Molecule type: protein  
A:Residues: 658-685; 689-696; 893-901 <TA2>  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleotide binding; nucleus; P-loo  
F:21-51/Region: zinc finger  
F:128-165/Region: zinc finger  
F:200-220/Region: helix-turn-helix motif  
F:224-231/Region: nuclear location signal  
F:250-270/Region: helix-turn-helix motif  
F:494-501/Region: nucleotide-binding motif A (P-loop)  
F:890-903/Region: nucleotide binding #status predicted

Query Match 25.3%; Score 704.5; DB 1; Length 1016;  
Best Local Similarity 34.4%; Pred. No. 1.3e-42;  
Matches 187; Conservative 95; Mismatches 213; Indels 49; Gaps 20;

Qy	1	MAPKKASVQTEGSKKQKQGTEDSFRSTAE--ALRAAPADNRVIRVDPSPCFPSRNPQ	57
Db	495	VGPRKSGAAP--SKKSGPVKECTGNSEKRMKLTLAGGAA-----VDPDSGLEHNAH	546
Qy	58	I--QVHEDYDCTLNQTNIGNNNKFFIYIQLLEEG--SRFFCMNRWRGVRGEV--GQSKMNH	112
Db	547	VLEKGGKVFSAITGLVDIVKGTNSYKQLQLEDDKENRYWIFRSWGRVGVIGSNKLEQ	606
Qy	113	TCLEDAKKDFKKFWKTKNKEWRDRFVAQPNKYIILIEVOGAESEAVVK--VDSGPV	170
Db	607	PSKEDAIEHFMKLYEKTGNAMHSKN-FTKHPKRYPLEID-YGODEEAVVKLTVPNGTK	664
Qy	171	RTVVKPCLDPATQNLITNIFSKEMFNKNAWTLNLDVKKMPLGLKLTQKQIARGFEAL	230
Db	665	SKLPKP-----VQNLIKMIFDVESMKKAMVEYIDLQKMPGLKSKRQIQAAYSILSEV	718
Qy	231	BEAMKNPTDGGQSLSELSFCYTVIPHNFGRSPPPIPSDVLOAKKMDLLVLADIQLAA	290
Db	719	QQAISQGSDDSHIL-DLSNRFYTLIPHDFGMKPPILLNADSVQAKVEMLDLDDIEVAY	777
Qy	291	TLQAAPGSEEEKVEVPHPLDRIYQLLRQLQLDLSGSEYKAIQIYTLQKQ----TGN	346
Db	778	SLLRGSDSSK-----DPIDVNYEKLKTDIKVYDRDSEAEIIRKYVKNTHATTNAYD	832
Qy	347	CPNLRHWKVNREGRDFOAHSKLGNRRLLWHGTNVAVVAAILTSGLRIMPH----SGG	402
Db	833	L-EVVDIEFKIEREGESQRYKPFQKLNRRLLWHGSRRTTNFAGILSQGLRIAPPEAPVTG	891
Qy	403	RVGKIYFASENSKSAGYVYTMHCGHGVYMFGLGEVALGKEHHITIDPSPKSPPPG	462
Db	892	MFGKIYFADMVSKSANYCHTSQ--GDPIGLILLGEVALGNMYELK-HARHISKLPKGRH	948
Qy	463	SVIARGQTEPDAQDIELELDGQPVVPPQVPPVQCFPSKSSFSQSEYLIYKESQCLRY	522
Db	949	SVKGLGKTPDPSASI--TVDG--VEVPLGTGIS-SCVNDTCLLYNEYIVYDIAQVHLK	1003
Qy	523	LLEI 526	
Db	1004	LLKL 1007	

RESULT 5  
T01311  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - Arabidopsis thaliana  
N:Alternate names: poly(ADP-ribose) polymerase; protein T14P8.19  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 21-Jul-2000  
C:Accession: T01311; S65662  
R:Kalicki, J.; Elliott, G.; Cloud, J.



A:Reference number: S04200; MUID:89263780

A:Accession: S04200

A:Molecule type: DNA

A:Residues: 1-1013 <HUP>

A:Cross-references: EMBL:X14206; NID:q49893; PIDN:CAA32421.1; PID:g49894

C:Genetics:

A:Map position: 1

C:Superfamily: NAD+ ADP-ribosyltransferase

C:Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc finger

Query Match 24.7%; Score 688; DB 1; Length 1013;  
Best Local Similarity 32.7%; Pred. No. 2e-41;  
Matches 177; Conservative 106; Mismatches 214; Indels 44; Gaps 17;

Qy 1 MAPKKASVQTEGSKQKQGTTEEEDSFSTAEALRAAPADNRVIRVDPSCFPSPRPGTQV 60  
Db 493 VAPRGSAAPSKSK-----GCFKEGVNKSSEKMKLTLLKGAAVDPSGLGSHSAHVLEKG 548  
Qy 61 HEDYDCTLNQTNIGNNNKFYIIQLLEG--SRFFCWNWRGVRGEV--GQSKMNHFTCLDED 117  
Db 549 GKVFATLGLVDIVKGTNSYKQLLEDKESRYWIFRSWGLGTIVIGSNKLEQMPSKEE 608  
Qy 118 AKKDPKKFKWEKTKNKEERDRFVAQPNKYTLIEVOGAESEAVVQVDSGPVRTVVKP- 176  
Db 609 AVEQPMKLYEEKTGNMWSKN-FTKYPKKFYPLEID-YGODEEAVVKL-----TVKPG 659  
Qy 177 --CSLDPATONLITNIFSKEMFKNMTLMNLDVKKMPLGKLTQKQIARGFEALEALEAM 234  
Db 660 TKSLLPKPVQELVGMIFDVDSMKKALVEYIDLQKMPGLKLSRRQIAQYAILSEVQQPV 719  
Qy 235 KNPTGDGQSLSELSCTFYVTPHNFGRSRPPPIINSPDVLOAKMDMLLVLADIELAOTLOA 294  
Db 720 SQGSSESQIL-DLSNRFYTLPHDGMKKPPLNNADSVQAKVEMLDNLLDIEVAYSLLR 778  
Qy 295 APGEEKEVEVPHPLDRDYQLLRCQLQDLSGESEYKAIQTYLKQ-----TGNISYRCNPL 350  
Db 779 GGSDDSSK-----DPIDVNYEKLTKDKVDRDSEAEVIRKYVKNTHATTNAYDL-EV 832  
Qy 351 RHVKNREGSGDRFOAHSKLGNRRLWHGNTNVAVVAAILTSGLRIMPH-----SGRVRGK 406  
Db 833 IDIFKIEREGESQRYKPPRLQHNRRLLWHGSRRTNFAGILSGURLIAPPEAPVTGYMFGK 892  
Qy 407 GIYFASNSKAGYVTTMHCQGHGVYMFGLGEVALGKHHITIDDPKSLKPPPGFDSVIA 466  
Db 893 GIYFADWYKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKCKHVKG 949  
Qy 467 RGQTEPDPAQDIELDGQPVVPGPVQCPSP-FKSSFSQSEYLYIKESQCRRLRYLLE 525  
Db 950 LGKTTDPDSASITLLE---GVEVPLGTGI--PSGVNDTALLYNEVYVYDIAQVNLKYLK 1003  
Qy 526 I 526  
Db 1004 L 1004

RESULT 8

S26057

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - rat (fragment)

C:Species: Rattus norvegicus (Norway rat)

C:Date: 11-Mar-1998 #sequence\_revision 17-Apr-1998 #text\_change 18-Jun-1999

C:Accession: S26057; S78453; I52331

R:Thibodeau, J.; Gradwohl, G.; Dumas, C.; Clairoux-Moreau, S.; Brunet, G.; Penning, C.;

Biochem. Cell Biol. 67, 653-660, 1989

A:Title: Cloning of rodent cDNA encoding the poly(ADP-ribose) polymerase catalytic domain

A:Reference number: I52331; MUID:90027702

A:Accession: S26057

A:Molecule type: mRNA

A:Residues: 1-500 <TH>

A:Cross-references: EMBL:X65497; NID:g56849; PIDN:CAA46478.1; PID:g56850

R:Potvin, F.

submitted to the EMBL Data Library, March 1992

A:Reference number: S78453

A:Accession: S78453

A:Molecule type: mRNA

A:Residues: 1-124 'H', 126-127, 'A', 129-238, 'D', 240-500 <POT>

A:Cross-references: EMBL:X65497

C:Superfamily: NAD+ ADP-ribosyltransferase

C:Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase

Query Match 24.6%; Score 686; DB 2; Length 500;  
Best Local Similarity 36.2%; Pred. No. 9.9e-42;  
Matches 173; Conservative 86; Mismatches 179; Indels 40; Gaps 17;

Qy 64 YDCTLNQTNIGNNNKFYIIQLLE--EGSRFFCWNWRGVRGEV--GQSKMNHFTCLDED 120  
Db 39 FSATLGLVDIVKGTNSYKQLLEDKESRYWIFRSWGLGTIVIGSNKLEQMPSKDAVE 98  
Qy 121 DFKKFKWEKTKNKEERDRFVAQPNKYTLIEVOGAESEAVVQVDSGPVRTVVKP---C 177  
Db 99 HFMKLYEEKTGNMWSKN-FTKYPKKFYPLEID-YGODEEAVVKL-----AVKPGTKS 149  
Qy 178 SLDPATONLITNIFSKEMFKNMTLMNLDVKKMPLGKLTQKQIARGFEALEALEAMKNP 237  
Db 150 KLPRPVQELVGMIFDVDSMKKALVEYIDLQKMPGLKLSRRQIAQYAILSEVQQVSG 209  
Qy 238 TGDGQSLSELSCTFYVTPHNFGRSRPPPIINSPDVLOAKMDMLLVLADIELAQTLOAAPG 297  
Db 210 SSESQIL-DLSNRFYTLPHDGMKKPPLNNADSVQAKVEMLDNLLDIEVAYSLLRGS 268  
Qy 298 EEEKVEEVEVPHPLDRDYQLLRCQLQDLSGESEYKAIQTYLKQ-----TGNISYRCNLRHV 353  
Db 269 DSSK-----DPIDVNYEKLTKDKVDRDSEAEVIRKYVKNTHATTNAYDL-EVIDI 322  
Qy 354 WKVNRREGSGDRFOAHSKLGNRRLWHGNTNVAVVAAILTSGLRIMPH-----SGRVRGKGIY 409  
Db 323 FKIEREGESQRYKPPRLQHNRRLLWHGSRRTNFAGILSGURLIAPPEAPVTGYMFGKGIY 382  
Qy 410 FASENSKAGYVTTMHCQGHGVYMFGLGEVALGKHHITIDDPKSLKPPPGFDSVIARGQ 469  
Db 383 FADWYKSANYCHTSQ--GDPIGLIMLGEVALGNMYELK-HASHISKLPKCKHVKGLGK 439  
Qy 470 TEPDPAQDIELDGQPVVPGPVQCPSP-FKSSFSQSEYLYIKESQCRRLRYLLEI 526  
Db 440 TAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTALLYNEVYVYDIAQVNLKYLK 491

RESULT 9

S31735

NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - African clawed frog (fragment)

N:Alternate names: poly ADP-ribose polymerase

C:Species: Xenopus laevis (African clawed frog)

C:Date: 22-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 20-Jun-2000

C:Accession: S31735; PN0495

R:Sautter-le Brean, B.M.

submitted to the EMBL Data Library, May 1992

A:Reference number: S31735

A:Accession: S31735

A:Molecule type: mRNA

A:Residues: 1-998 <SAU>

A:Cross-references: EMBL:Z12139; NID:g64967; PIDN:CAA78126.1; PID:g1334661

R:Ozawa, Y.; Uchida, K.; Uchida, M.; Ami, Y.; Kushida, S.; Okada, N.; Miwa, M.

Biochem. Biophys. Res. Commun. 193, 119-125, 1993

A:Title: Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose) polymerase

A:Reference number: PN0494; MUID:93277538

A:Accession: PN0495

A:Molecule type: mRNA

A:Residues: 742-745, 'E', 747-876 <OZA>

C:Comment: This zinc-finger protein plays a role in DNA repair, cell growth, and diff

C:Superfamily: NAD+ ADP-ribosyltransferase

C:Keywords: DNA binding; glycosyltransferase; hexosyltransferase; NAD; nucleus; pento

Query Match

Best Local Similarity 24.4%; Score 681; DB 2; Length 998;

A:Reference number: S31735; Pred. No. 6.1e-41;

Matches 178; Conservative 95; Mismatches 215; Indels 42; Gaps 18;

```

Qy 10 QTEGSKOROGT--EEDSFRSTAEALRAAPADNRVIRVDSPCFPSRNPBGIOVHVEDCTL 68
Db      : : | : : : : : : : : : : : : | : : : : : : : : : : : :
488 KSSKVREEGKSEKKMKLTVKGAADPDS---ELEDSCHVLETTG---KIFSATL 540
Db      : : | : : : : : : : : : : : : | : : : : : : : : : : : :
Qy 69 NQTWIGNNNKFYIOLLE--EGSRFFCMNWRGRVGVEV-QOSKMNHFTCLDEDAKKFKKK 125
Db      : : | : : : : : : : : : : : : | : : : : : : : : : : : :
541 GLVDITRGTSYYKLQILIEHDRSDRYWFRSGRWGVTVIGSKKLEEMSSKEDAIEHFLNL 600
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
Qy 126 FWEKTKNKWEERDRFVAQPNKYLLIEVQ-CEARSQEAVKVDGSPVRTVVYPCSLDPATO 184
Db      : : | : : : : : : : : : : : : | : : : : : : : : : : : :
601 YQDKRTGNAWHS-PNFTKYPKKFPYLEIDYGQEDDVKKLSVGAGTKSKLAKP-----VQ 653
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
185 NLINIFESKMFKNAMTLMLNDLVKKMPLGLKTQQIARGFEALEEAMKNPTGGQSIL 244
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
654 ELIKLIFDVESMMKAMVEFIDLQKMPGLKSKRKIQGSAYSILSQVQAQVSESILSEARLL 713
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
245 EELSCTCYTVIPHNFGRSRPPPIINSPOVLQAKMDMLLVADIELAQTLOAAPGEKEKVE 304
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
714 -DLSNQFYITLPHDFGMKKPKPLLNNLEYIOAKVQMLDNLLDIEVAYSSLRGAGDDGEC-- 770
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
305 EVPHLPDRDYOLLRCQLQLLDGSEYKAOTYLK-----QTGNSYRCPNLRHVMKYVNREG 360
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
771 ---DPIDVKEYIKTDIKVYAKOSESRIICDYVKNTHADTHNAYDLEVLF-EJFKIDREG 826
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
361 EGDRFQASHKLGNNRLLNHGHTNVAVAAIITSGLRIMPH----SGGRVKGVIYFASENSK 416
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
827 EYQRKYKPFQJLHNRLQHLWHGSRITTFNAGILSOGLRIAPPAPVTGYMGKGIYFADMVSK 886
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
417 SAGVYTMHCGGHGOVGYMFLGEVALGHEHITTDDPSLKSPPGCFDSVIARTGETPDPAQ 476
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
887 SANTYCHAM--PGSPIGLILLIGEVALGNHNEHKLKAASQITKL-PKGKHSVKGLGRTADPDSA 943
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
Qy 477 DIELELDGPVVVPQGPPVQCPSFKSSFSOSEYLIYKESQCRRLYILLEI 526
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :
944 TV--QJLDG--VDVPLRGKGTSA-NISDFTSLLYNEYIVYDIQAQVNLKYLLKL 998
Db      : : | : : : : : : : : : : : : | : : : : ~~~~~~ : : : :

```

RESULT	10
--------	----

S42208  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30) - flesh fly (*Sarcophaga peregrina*)  
N:Alternate names: poly(ADP-ribose) polymerase  
C:Species: *Sarcophaga peregrina*  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: S42208; S71496  
E:Masutani, M.; Nozaki, T.; Hitomi, Y.; Ikejima, M.; Nagasaki, K.; de Prati, A.C.; Kurat  
Eur. J. Biochem. 220, 607-614, 1994  
A:Title: Cloning and functional expression of poly(ADP-ribose) polymerase cDNA from *Sarcophaga*  
A:Reference number: S42208; MUID:941170813  
A:Accession: S42208  
A:Molecule type: mRNA  
A:Residues: 1-996 <MAX>  
A:Cross-references: EMBL:D16482; NID:g473742; PIDN:BAA03943.1; PID:g538248  
A:Accession: S71496  
A:Molecule type: protein  
A:Residues: 170-188;721-736;813-819;879-885 <MAX>  
A:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase; zinc finger  
F:1-369/Domain: DNA binding #status predicted <DNA>  
F:370-507/Domain: auto-modification #status predicted <AMO>  
F:508-996/Domain: NAD binding #status predicted <NAD>

Query Match 23.4%; Score 652; DB 1; Length 996;  
Best Local Similarity 33.3%; Pred. NO. 7.6e-39;  
Matches 176; Conservative 100; Mismatches 205; Indels 48; Gaps

Qy	22	EEEDSFSTAEALRAAPADNRV	----	IRVDPSCP	-----	SRNPGIVQHYEDYDCTL	68
		:  :  :  :  :  :				:  :	
Db	489	EESSKSSKSYLYTKSPKSWTLK	IKGLA	VDPSGLDVAHVYVSRN	-----	KERYNVVL	543
Qy	69	NOTNIGNNNNKFYIIQLLESG	--	SRPFQWNRWGRV	-	EVGQSKMNFHFTCLEDAKKDFKK	125

	I : I   I   I   I   I   I   I   I   I   I   I   I   I   I   I   I   I   I	
Db	544 GTIDIQKKNKSFYKLQLESMDKNRWFPSRGRTGTTGGHKNLFNSLWDAIYQFKEL	603
Qy	126 FWEKTNNKWBEDRFQAQPNKYTLIEVGEABEQEAQVVDSPGVTVTKPCLSDPATQN	185
Db	604 YLEKSGHNFRNFVKYAGRWPIDIDYAEDS----KIDLSEAHDIKSLPL--SVQD	656
Qy	186 LITNIFSKEMFKNAMTLMNLVDKKMPLGLTKTOQTARGFEALEALEEAMKNPTDGQSLE	245
Db	657 IIKLMPDVDMSKTRMEFDLDMEKPLGLSQIQOISAYKVLTETVELIQ--GGTNAKFI	715
Qy	246 ELSDCFVTIVPHNFGSRRPINSPDVLOAKKDMLLVLADIELAQTLQNAPEEEKEVEE	305
Db	716 DATNRFTYLPHNFGTSQPPLDTTEQVEQLROMLDSLIEICAYSLLQT---EDSKAD-	771
Qy	306 VPHPLEDROYQLLCOLQLDSDSESEYKATOTYLKQT-GNSYRCPNLR-HVMXVRNEGEG	362
Db	772 -INPDKHYBQLTKLEPLDNKNEEYILQYVKNTHAETHLYLDEVVDIFKVARQGEA	830
Qy	363 DRFOAHSKLGNRRLLMHGTNVAVVAAILTSGLRIMPH-----SGGRVGKGIIYFASENSKA	418
Db	831 RRYAPFKLNHRLLMHGSRLLTNFAGILSHGLKIADPEAPVTGYMGKGIIYFADMVSKA	890
Qy	419 GYVTTMHCGGHQGYMFLGEVALGCKEHHTIIDDPSLKSPPGFDSVIARQGETPDPAODI	478
Db	891 NYCCTSH--HNSGTMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFCGRGMTMPNPSEI	947
Qy	479 ELELDGQGVVVPPGPVCQPSFKSSFSOSEYLIYKESOCRLRYLLEIH	527
Db	948 IRE-DG--VEIPLGKPTINDSLK-SLLNYNEFTIYDIAOWNTQYMLRMN	992

RESULT 11

T51353  
NAD+ ADP-riboseyltransferase [EC 2.4.2.30] [imported] - Arabidopsis thaliana  
N:Alternate names: poly(ADP-ribose) polymerase  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 18-Aug-2000 #sequence\_revision 18-Aug-2000 #text\_change 02-Sep-2000  
C:Accession: T51353  
R:Doucet-chabeaud, G.; Kazmaier, M.  
submitted to the EMBL Data Library, December 1998  
A:Reference number: Z25379  
A:Accession: T51353  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: mRNA  
A:Residues: 1-983 <DOU>  
A:Cross-references: EMBL:AJ131705; PIDN:CAAL0482.1  
A:Experimental source: cultivar landsberg erecta  
C:Genetics:  
A:Gene: parp-1  
C:Function:  
A:Description: ADP-ribose polymer synthesis  
C:Superfamily: NAD+ ADP-ribosyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; pentosyltransferase

Query Match 22.4%; Score 624; DB 2; Length 983;  
Best Local Similarity 33.0%; Pred. No. 7.8e-37;  
Matches 179; Conservative 87; Mismatches 209; Indels 68; Gaps 21;

Qy	15	KKQRCGTTEEDSFRSTAALRAAPADNRVIRVDPSCFSPSRNPGLQVH----	ED----	YDC	60
Db	472	KKQRLPDKYKIEDTSLSLTVKVKGRSAVHEAS-----GLQEHCHILEDGNISYNT			524
Qy	67	TLNQTNGNNNNKYYIIQLLEE--GSFFCWNRWGRVG--EVGSGKNHHTCLEDKAKDF			122
Db	525	TLSMSDLSTGINSYIIQLITQEDKSGDCYFVRKWRGVNEKIGGNKVEEMS-KSDAVHEF			583
Qy	123	KKKFEWTKN--KWEERDRFAQPNKYTLILEYQGEAESQEAIVVVDSGGVRTVWPKCSL			179
Db	584	KRLFELEKTNWSEWSEQKTNFQKQPGKFLDLIDYGVNKKQVA-----KKPEFQT----			636
Qy	180	DPATONLITNFSKMEKFNKAMTLMLNDVKKKMLPKGLTKKOOIARGFEAL--			235

Db 637 APSLIELMKMLFDVETYSAMMEFEINNSEMPGLKSLKHNQKGFALTEIQRLLTESDP 696  
QY 236 NPTGQOSLEELSSCFYVIPHNFGRSRPPPIINSDPVLOAKKMDLLVLADIQAOTLQAA 295  
Db 697 OPTMKESLLVDASNRFFTMIP-----SIHPHIIRDEDDFKSKVKMLEALQDIEIASRI--- 749  
QY 296 PCEEEKVEEVPHPDLDRDYQLLRCOLOLLDGSSEYKAIQTYLKTGNSYRCP----- 348  
Db 750 VGFDVSTES-----LDDYKKLHCDISPLPHDSEYRIEIKYL-----NTHAPTHTESWL 801  
QY 349 NLRHWKYNRREGEDRFOAH-SKLGNNRLLMHGTNVAVVAAILTSLGRI-----MPHSGGR 403  
Db 802 ELEEVFALEREGEFKYPAPHREKLGKMLLWHGSLTNFVGLNQLGRLIAPPEATGYM 861  
QY 404 VKGIYFASSENSKAGYVTTMCHGGHGVGMFLGVALGKEHHITIDPSLSKSPPPGDS 463  
Db 862 FGKGIYFADLVSKSAQCYT--CKKNPVGLMLLSEVALGEIHEL-T-KAKYMDKPPRGKHS 918  
QY 464 VIARQTEPDPAQDIELELDGQPVVPOGPPVQCPSPFKSSFSQSEYLIYKESQCLRYL 523  
Db 919 TKGLGKKVP---QDSEFAKWRGDTVPCGKPV-SKVRKASELMYNEYIYDTAQVKLQFL 974  
QY 524 LEI 526  
Db 975 LKV 977  
RESULT 12  
C84719  
probable poly (ADP-ribose) polymerase [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: C84719  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, A.  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: A84420; MUID:20083487  
A:Accession: C84719  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1009 <STO>  
A:Cross-references: GB:AE002093; NID:5432827; PIDN:AAD20677.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g31320  
A:Map position: 2  
C:Superfamily: NAD+ ADP-riboseyltransferase  
Query Match 22.4%; Score 624; DB 2; Length 1009;  
Best Local Similarity 33.08; Pred. No. 8.1e-37;  
Matches 179; Conservative 87; Mismatches 209; Indels 68; Gaps 21;  
QY 15 KKQGTGTEEDSFRSTAEALRAAPADNRVIRVDPSPCFSRNPGIQVH-----ED-----YDC 66  
Db 498 KKQRLPDKYKIDTESLTVKVKGRSAVHEAS-----GLQEHCHILEDGNSIYNT 550  
QY 67 TLNQTNNNNKFFIYIQLLEB--GSRFCWNRWGRVG--EVGOSKMHFTCLEDAKDKDF 122  
Db 551 TLSMSDLSTGINSYVLIQIEDKGSDCYVFRKWRGVRGNEKIGGNKVEEMS-KSDAVHEF 609  
QY 123 KKKWEKTKN--KWEEDRFRVAQPNKYTLIEVOGEAESQEAQVNVKVDGSPVTVVKPCSL 179  
Db 610 KRLFEKTGNTWSEQKTFKQKPGKPLDIDYGVNKQVA-----KKEPFQT---SSNL 662  
QY 180 DPATONLTNIFSKEMFKNAMTLMNLDVKMKPLGLTKQQLARGFEAL-----EALAEAMK 235  
Db 663 APSLIELMKMLFDVETYSAMMEFEINNSEMPGLKSLKHNQKGFALTEIQRLLTESDP 722  
QY 236 NPTGQOSLEELSSCFYVIPHNFGRSRPPPIINSDPVLOAKKMDLLVLADIQAOTLQAA 295

Db 723 OPTMKESLLVDASNRFFTMIP-----SIHPHIIRDEDDFKSKVKMLEALQDIEIASRI--- 775  
QY 296 PCEEEKVEEVPHPDLDRDYQLLRCOLOLLDGSSEYKAIQTYLKTGNSYRCP----- 348  
Db 776 VGFDVSTES-----LDDYKKLHCDISPLPHDSEYRIEIKYL-----NTHAPTHTESWL 827  
QY 349 NLRHWKYNRREGEDRFOAH-SKLGNNRLLMHGTNVAVVAAILTSLGRI-----MPHSGGR 403  
Db 828 ELEEVFALEREGEFKYPAPHREKLGKMLLWHGSLTNFVGLNQLGRLIAPPEATGYM 887  
QY 404 VKGIYFASSENSKAGYVTTMCHGGHGVGMFLGVALGKEHHITIDPSLSKSPPPGDS 463  
Db 888 FGKGIYFADLVSKSAQCYT--CKKNPVGLMLLSEVALGEIHEL-T-KAKYMDKPPRGKHS 944  
QY 464 VIARQTEPDPAQDIELELDGQPVVPOGPPVQCPSPFKSSFSQSEYLIYKESQCLRYL 523  
Db 945 TKGLGKKVP---QDSEFAKWRGDTVPCGKPV-SKVRKASELMYNEYIYDTAQVKLQFL 1000  
QY 524 LEI 526  
Db 1001 LKV 1003  
RESULT 13  
A47474  
NAD+ ADP-riboseyltransferase (EC 2.4.2.30) - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: A47474  
R:Uchida, K.; Hanai, S.; Ishikawa, K.; Ozawa, Y.; Uchida, M.; Sugimura, T.; Miwa, M.  
Proc. Natl. Acad. Sci. U.S.A. 90, 3481-3485, 1993  
A:Title: Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase: leucine zip  
A:Reference number: A47474; MUID:93234521  
A:Accession: A47474  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-994 <UCH>  
A:Cross-references: GB:D13806; GB:D13807; GB:D13808; NID:303545; PIDN:BA002964.1; PI  
A:Note: sequence extracted from NCBI backbone (NCBIN:129703, NCBIp:129704)  
C:Genetics:  
A:Gene: FlyBase:Parp  
A:Cross-references: FlyBase:FBgn0010247  
C:Superfamily: NAD+ ADP-riboseyltransferase  
C:Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase  
Query Match 22.3%; Score 621.5; DB 1; Length 994;  
Best Local Similarity 31.6%; Pred. No. 1.2e-36;  
Matches 167; Conservative 102; Mismatches 221; Indels 39; Gaps 17;  
QY 16 KORQGTGTEEDSFRSTAEALRAAPADNRVIRVDPSPCFSRNPGIQVH-----VHED-----YDCT 67  
Db 481 KSRPKETTKSLNSNIYTKSPV-SRTFKVKDGLAVDPDGLIEDIAHYVDNSNNKYSV 539  
QY 68 LNQTNNNNKFFIYIQLL--BEGSRFCWNRWGRVG--EVGOSKMHFTCLEDAKDKFKK 124  
Db 540 LGLTDIQRNKNSYKVLKADKKEKYWIFRSWGRIGTNIGNSKLEEFDTSETSAKRNFKE 599  
QY 125 KFEWTKNKKWEEDRFRVAQPNKYTLIEVOGEAESQEAQVNVKVDGSPVTVVKPCSLDPATQ 184  
Db 600 IYADKTGNEYEQDRNFVKTRGRMYPETIQ--YDDQKLVKKHSHFTS-----KLEISVQ 652  
QY 185 NLITNIFSKEMFKNAMTLMNLDVKMKPLGLTKQQLARGFEALAEAMKNPTGDGSL 244  
Db 653 NLKILIFIDISMNKTLMEFHIDMKPLGLKLSAHOIQSAIRVVKIYVNWLEGGSTAK-L 711  
QY 245 EELSSCFYVIPHNFGRSRPPPIINSDPVLOAKKMDLLVLADIQAOTLQAAPEEGEEKVE 304  
Db 712 IDATNRFYTLIPHNEGVOLPTLIETHQOIEDLRQMLDSLAEIYVAYS-----IKSEDSV 766  
QY 305 EYPHPDLDRDYQLLRCOLOLLDGSSEYKAIQTYLKTGNSYRCPNLR--HWKYNRGE 361  
Db 767 DACNPLDNHYAQIKTQLVALDKNSEEFSLSQYVKNTHASTHKSVDLKVDFVFKYSRQGE 826





**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:24 ; Search time 36.19 Seconds

(without alignments)  
564.905 Million cell updates/sec

Title: US-09-701-586B-10

Perfect score: 2789

Sequence: 1 MAPKKASVQTEGSKRQGG.....EVLIIKESQCLRYLLEIHL 528

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2241.5	80.4	533	1 PPO3_HUMAN	Q9y6f1 homo sapien
2	718	25.7	1012	1 PPO2_CRIGR	Q9r152 cricetus
3	712.5	25.5	559	1 PPO2_MOUSE	O88554 mus musculus
4	707.5	25.4	1013	1 PPO2_HUMAN	P09874 homo sapien
5	704.5	25.3	1015	1 PPO2_BOVIN	P18493 bos taurus
6	699	25.1	583	1 PPO2_HUMAN	Q9ugn5 homo sapien
7	698	25.0	1013	1 PPO2_RAT	P27008 rattus norv
8	697.5	25.0	637	1 PPO2_ARATH	Q11207 arabidopsis
9	694	24.9	1011	1 PPO2_CHICK	P26446 gallus gall
10	688	24.7	1012	1 PPO2_MOUSE	P11103 mus musculus
11	681	24.4	998	1 PPO2_XENLA	P31669 xenopus lae
12	652	23.4	996	1 PPO2_SARPE	Q11208 sarcophaga
13	621.5	22.3	994	1 PPO2_DROME	P35875 drosophila
14	389	13.9	538	1 YQ4_CAEEL	Q09525 caenorhabdi
15	276.5	9.9	1724	1 PPOV_HUMAN	Q9uk33 homo sapien
16	193	6.9	135	1 PPO2_ONCMA	Q08824 oncorhynch
17	113.5	4.1	2035	1 EVPL_MOUSE	Q9d952 mus musculus
18	107.5	3.9	638	1 TOXA_PSEAE	P11439 pseudomonas
19	106.5	3.8	646	1 YE14_SCHPO	O13869 schizosacch
20	105.5	3.8	3075	1 LMAL_MOUSE	P25391 homo sapien
21	103	3.7	379	1 Y01L_MOUSE	P11260 mus musculus
22	103	3.7	1031	1 KINH_STRPU	P35978 strongyloce
23	102.5	3.7	612	1 EXO2_BPT5	P11109 bacterioph
24	101	3.6	1574	1 RPOC_AQUAE	O67763 aquifex aeo
25	100.5	3.6	629	1 PAB2_ARATH	P42731 arabidopsis
26	100.5	3.6	830	1 YEMA_SCHPO	Q10332 schizosacch
27	100.5	3.6	890	1 RB6K_HUMAN	O95235 homo sapien
28	100.5	3.6	1395	1 SP4L_YEAST	P38904 saccharmyc
29	100.5	3.6	3210	1 GENE_HUMAN	P49454 homo sapien
30	100	3.6	2025	1 TTC3_HUMAN	P53804 homo sapien
31	99	3.5	583	1 T2F1_FLAOK	P14870 flavobacter
32	98	3.5	2453	1 NCRL_MOUSE	Q60974 mus musculus
33	97.5	3.5	548	1 SYFB_AERPE	Q9y913 aeropyrum p

34	97	3.5	727	1 MPPI_ARATH	Q91w85 arabidopsis
35	96.5	3.5	495	1 NUSA_ECOLI	P03003 escherichia
36	96.5	3.5	497	1 BIR4_HUMAN	P98170 homo sapien
37	96.5	3.5	583	1 LAM1_XENLA	P09010 xenopus lae
38	96.5	3.5	808	1 PML_MOUSE	Q60953 mus musculus
39	96.5	3.5	853	1 PBPA_HAEIN	P31776 h penicilli
40	96.5	3.5	1396	1 RPOC_CHLMU	O9pk79 chlamydia m
41	96.5	3.5	3587	1 SRF2_BAGSU	Q04747 bacillus su
42	96	3.4	605	1 VCLA_GOSHI	P05799 gossypium h
43	96	3.4	963	1 KINH_HUMAN	P33176 homo sapien
44	96	3.4	1875	1 KLPI_YEAST	Q02455 saccharomyc
45	95.5	3.4	725	1 ADDB_MOUSE	Q9qyb8 mus musculus

#### ALIGNMENTS

RESULT 1

PO3\_HUMAN  
ID PPO3\_HUMAN STANDARD; PRT; 533 AA.  
AC Q9Y6F1: Q9UG81;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-riboyltransferase-3) (Poly[ADP-ribose] synthetase-3) (PADPRT-3) (hPARP-3).  
DE (hPARP-3).  
GN ADPRTL3 OR PARP3 OR ADPRT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=99263509; PubMed=10329013;  
RA Johansson M.;  
RT "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA cloning of two novel poly(ADP-ribose) polymerase homologues.";  
RL Genomics 57:442-445(1999).  
[2]  
SEQUENCE OF 75-533 FROM N.A.  
RN TISSUE-Kidney;  
RC Ansgore W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.; Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RL -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor = nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.  
CC -1- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -1- TISSUE SPECIFICITY: Widely expressed: the highest levels are in the kidney, skeletal muscle, liver, heart and spleen; also detected in pancreas, lung, placenta, brain, leukocytes, colon, small intestine, ovary, testis, prostate and thymus.  
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

EMBL; AF083068; AAC29855.1; -;  
DR EMBL; AL050034; CAB43246.1; -;  
DR HSSP; P26446; 1A26.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Transferase; Glycosyltransferase; NAD; Nuclear protein;  
KW ADP-riboseylation.  
FT DOMAIN 14 20 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT CONFLICT 80 80 K -> N (IN REF. 2).



```
Db 493 APKGSAAPSKSKS---GLYKEGVNKKRMLTKGGAADVDPDGLHSAHVLEKGG 548
QY 62 EYDCITLQNTNIGNNNFYIIQLLEGG--SRFFCNRWRGRVGEV-GOSKMMHFTCLEDA 118
Db 549 KYFSATLGLVDIVKGTNSYKYLQLEDDKESYWIWFRSGRVTGTVIGSNKLEQMPKSKEDA 608
QY 119 KDKFKKFKWEKTKNKEWRDRVAQPNKVTLLTEVOGEAESQAVKVDSPGVRTVVKP-- 176
Db 609 VEHFKMLYEKTKNAHNSKN-FTKYPKFFPLEID-YGODEAVVKL-----TVKPGT 659
QY 177 -CSLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGLKTKQQTARGFEALEALEEAMK 235
Db 660 KSKLPKAVQELVGMIFVDSMKKALVEYEDIQKMPGLGSKRQIQAAVSYLSEVQOAVS 719
QY 236 NPTGQSLLEELSSCFYVIVPHNFRSRRPPPPINPDVQAKKMDMLLVADIETLAOTLOA 295
Db 720 QGSSDSQIL-DLSNRYFTLPHDFGKMKPPPLNNADSVQAKYEMDLNLDIEVAYSLRG 778
QY 296 PEEEEKVEEVPHPDRDYQLRLCQLQLLSDGSEYKATQYTKQ---TGSNYRCPNLR 351
Db 779 GSDSSK-----DPIDVNYEKLTKDKVDRDSEAEVIRKVKWTHATHNAYDL-EVM 832
QY 352 HWKYNVREGDRFOAHKSLGNRRLLHGTNVAVVAALITSGLRIMPH-----SGGRVKGK 407
Db 833 DIFKIEREGESORYKPKQLHNRLLHGRSRTTNFAGILSQGLRIAPPEAVPTGYMFGK 892
QY 408 IFYFSENSKAGYVTVMCGGHGVGMFLGEVALGKEHHITIDPSLSPPPGFSVIA 467
Db 893 IFYFADWKSANYSCTSO--GDPIGLIILGEVALGNMYELK-HASHISKLPRKGRSVKGL 949
QY 468 GTEPDPAQDIELELDGQPVVPPQPPVQCPSP-FKSSFSQSEYLYIKESQCLRLVLEI 526
Db 950 GKTTFDPSASITL-----GVEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNLKYLKL 1003

RESULT 3
PPO2_MOUSE
ID PPO2_MOUSE STANDARD; PRT; 559 AA.
AC O88554; O39N29;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Poly (ADP-ribose) polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2)
DE (mPARP-2)
DE ADPRTL2 OR PARP2 OR ADPRT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Embryo;
RX MEDLINE=99292755; PubMed=10364231;
RA Ame J.-C., Rolli V., Nierdergang C., Apicou F., Decker P.,
RA Muller S., Hoger T., Schreiber V., de Murcia J., de Murcia G.M.;
RA "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)
RT polymerase.";
RL J. Biol. Chem. 274:17860-17868(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV;
RX MEDLINE=21179160; PubMed=11133988;
RA Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,
RA Nierdergang C.P.;
RT "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2
RT (PARP-2) gene to the gene for RNase P RNA.";
RL J. Biol. Chem. 276:11092-11099(2001).
RN [3]
RP SEQUENCE OF 9-559 FROM N.A.
RC STRAIN=129/Sv x C57BL/6;
RX MEDLINE=99268466; PubMed=10338144;
RA Berghammer H., Ebner M., Marksteiner R., Auer B.;
```

```
RT "PADPRT-2: a novel mammalian polymerizing (ADP-ribose) transferase gene
RT related to truncated PADPRT homologues in plants and Caenorhabditis
RT elegans.";
RL FEBS Lett. 449:259-263(1999).
CC -!- FUNCTION: HAS DNA-DEPENDENT POLY (ADP-RIBOSE) POLYMERASE ACTIVITY.
CC SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribose)(N)-acceptor -
CC nicotinamide + (ADP-D-ribose)(N+1)-acceptor.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in
CC testis followed by ovary.
CC -!- INDUCTION: By high levels of DNA-damaging agents.
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
EMBL; AJ007780; CAA07679.1; -.
EMBL; AF191547; AAK13253.1; -.
EMBL; AF072521; AAC25415.1; ALT_INIT.
DR HSSP; P26446; 1A26.
DR MGD; MGI:1341112; Adprt2.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP-reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP-reg; 1.
DR Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation.
FT DNA_BIND 1 65
FT DOMAIN 66 559
FT DOMAIN 3 9
FT DOMAIN 33 39
FT DOMAIN 82 82
FT CONFLICT 177 177 L -> V (IN REF. 2).
FT CONFLICT 486 486 V -> I (IN REF. 2).
FT CONFLICT 486 486 R -> Q (IN REF. 2).
SQ SEQUENCE 559 AA; 63396 MW; E0AEDAEE412C1445 CRC64;
```

Query Match 25.5%; Score 712.5; DB 1; Length 559;  
Best Local Similarity 35.2%; Pred. NO. 7.8e-45;  
Matches 204; Conservative 87; Mismatches 207; Indels 81; Gaps 25;

```
QY 1 MAPKKAS-----VQTEGSKQKQGTG---EEDS--FRSTAELRAAP-----ADN 41
Db 1 MAPRRQRSGSGRRVLNE-AKKVDNKNKATEDDSDPGKKMRTCORGPAGGKADRTKDN 59
QY 42 R-----VIRVDPSCPPSRNPG---IQVHEDYDCTLNQTNIGNNNKFFYIIQLLE 88
Db 60 RDSVTKTLKLGKAPVDPECAAKLGRAHYVCEGDDVDVMLNQTNLQFNKNNKYLIQLLED 119
QY 89 GSR--FFCWNWRGRVGEVQSKMNHFTC---LEDAKKDKPKKFKWEKTKNKEWRDRFVAQ 143
Db 120 DAQRNFSVMWRGRVGTGQHSLS--VTCSDGLNKAKEIFQKFKLDKTKNWNEDRENFEKV 177
QY 144 PNKYTLIEVQGEAESQ-EAAVVKVDSGVRTVVKPCS-LDPATQNLITNIFSKEMFNAMT 201
Db 178 PGKYDMLQNDYAAASTQDESKTKEE-----TLKPSQLDRVQVELKLKICNQTWEEMI 232
QY 202 LNNLDVKKMPLGLKTKQQTARGFEALEALEEAMKNPTGD-GOSLEELSSCFYVIVPHNFG 260
Db 233 EMKYDTKRAPLGLKTAQIKAGYQSLKKIEDCIR--AGQHGRALVEACNEFYTRIPHDG 290
QY 261 RSRPPPTNSPDVQAKKMDMLLVADIETLAQTLOAAGPEEEKVEEVPHPDRDYQLLRQC 320
Db 291 LSPVPVIRTEKELSDVKLLKLEALGDIETALKLKVS---ERQGLE---HPLDQHYRLHCA 344
```

```
QY 321 LQLDSGESEYKATQYLKOTGNSYRCP-----NLRHVWVKNVREGEGRFQAHSKLGN 373
Db 345 LRLPDHSENEFKVISOYLOST-----HAPTHKDYDTMTLIDVFEVEKEKEAFR--EDLPN 398
QY 374 RRLWHTGNVAVAAIILTSGLRIMPH---S GGRVKGKGIYFASENSKSAGVYTTMHCGGH 429
Db 399 RMLWHRGSRISNWWGILLSHGLRUPAPPAPITGYMFGGIYFADMSSKSANVCFA SRL--K 456
QY 430 QVGYMFLGCEVALGKREHITIDDPDLKSPPPGFDSDVIARGQTEPPPAQDIELELDGPVVV 489
Db 457 NTGLLLSEVALGOCNELLEANPRAQGLLRKHKSTKMGKMAPSPAIFI--TLNGS--TV 512
QY 490 PQGPVPQCPSEKSSSF--SQSEYLIYKESQRLRYLLEI 526
Db 513 PLGPASDTGILNPGGYTLNNEFTVYSPNQVMRYLLKI 551

RESULT 4
PPOL_HUMAN STANDARD; PRT; 1013 AA.
AC P09874;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE ADP-ribosyltransferase-1 (Poly[ADP-ribose] synthetase-1).
GN ADPRT OR PPOL OR PARP1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE-90091744; PubMed-2513174;
RA Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;
RT "Human nuclear NAD+ ADP-ribosyltransferase(polymenzing):
RT organization of the gene.";
RL DNA 8:575-580(1989).
[2]
RN SEQUENCE FROM N.A.
RX TISSUE-Fibroblast;
RX MEDLINE-88076933; PubMed-3120710;
RA Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,
RA Suzuki H., Nyunoya H., Miwa M., Sugimura T.;
RT "Nucleotide sequence of a full-length cDNA for human fibroblast
RT poly(ADP-ribose) polymerase.";
RL Biochem. Biophys. Res. Commun. 148:617-622(1987).
[3]
RN SEQUENCE FROM N.A.
RX TISSUE-Fibroblast;
RX MEDLINE-88058958; PubMed-2824474;
RA Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,
RA Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,
RA Inayama S., Shizuta Y.;
RT "Primary structure of human poly(ADP-ribose) synthetase as deduced
RT from cDNA sequence.";
RL J. Biol. Chem. 262:15990-15997(1987).
[4]
RN SEQUENCE FROM N.A.
RX MEDLINE-88068596; PubMed-2891139;
RA Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,
RA Hensley P., Smulson M.E.;
RT "cDNA sequence, protein structure, and chromosomal location of the
RT human gene for poly(ADP-ribose) polymerase.";
RL Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).
[5]
RN SEQUENCE OF 440-1013 FROM N.A.
RX MEDLINE-87298455; PubMed-3113420;
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA Miwa M.;
RT "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and
RT expression of its gene during HL-60 cell differentiation.";
RL Biochem. Biophys. Res. Commun. 146:403-409(1987).
```

```

RN [6]
RP ERRATUM.
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,
RA Miwa M.;
RL Biochem. Biophys. Res. Commun. 148:1549-1550(1987).
[7]
RN SEQUENCE OF 1-94 FROM N.A.
RX MEDLINE-91099327; PubMed-2125269;
RA Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,
RA Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,
RA Maeda T., Ikeda H., Sagara Y., Shizuta Y.;
RT "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter
RT region.";
RL Eur. J. Biochem. 194:521-526(1990).
[8]
RN SEQUENCE OF 1-39 FROM N.A.
RX MEDLINE-90211250; PubMed-2108670;
RA Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,
RA Esumi H.;
RT "Characterization of a putative promoter region of the human
RT poly(ADP-ribose) polymerase gene: structural similarity to that of
RL Biochem. Biophys. Res. Commun. 167:701-710(1990).
[9]
RN SEQUENCE OF 1-39 FROM N.A.
RA Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzer D.,
RA Schweiger M.;
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.
[10]
RN ANALYSIS OF ZINC FINGERS.
RX MEDLINE-90222155; PubMed-2109322;
RA Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,
RA Koken M.H.M., Hoeljmakers J.H.J., de Murcia G.M.;
RT "The second zinc-finger domain of poly(ADP-ribose) polymerase
RT determines specificity for single-stranded breaks in DNA.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).
[11]
RN ANALYSIS OF ZINC FINGERS.
RX MEDLINE-91072398; PubMed-2123876;
RA Ikellma M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,
RA Gill D.M., Miwa M.;
RT "The zinc fingers of human poly(ADP-ribose) polymerase are
RT differentially required for the recognition of DNA breaks and nicks
RT and the consequent enzyme activation. Other structures recognize
RL intact DNA.";
RL J. Biol. Chem. 265:21907-21913(1990).
[12]
RN MUTAGENESIS OF CATALYTIC DOMAIN.
RX MEDLINE-91035460; PubMed-2121735;
RA Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,
RA Molinete M., Penning C., Keith G., de Murcia G.M.;
RT "Expression and site-directed mutagenesis of the catalytic domain of
RT human poly(ADP-ribose)polymerase in Escherichia coli. Lysine 893 is
RT critical for activity.";
RL J. Biol. Chem. 265:19249-19256(1990).
[13]
RN NUCLEAR LOCALIZATION SIGNAL.
RX MEDLINE-92371433; PubMed-1505517;
RA Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,
RA Menissier de Murcia J.;
RT "The human poly(ADP-ribose) polymerase nuclear localization signal is
RT a bipartite element functionally separate from DNA binding and
RT catalytic activity.";
RL EMBO J. 11:3263-3269(1992).
[14]
RN MUTAGENESIS OF CATALYTIC DOMAIN.
RX MEDLINE-97461532; PubMed-9315851;
RA Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;
RT "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic
RT domain reveals amino acids involved in polymer branching.";
RL Biochemistry 36:12147-12154(1997)
CC -1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
```





RC TISSUE=Fetal brain;  
 RX MEDLINE=99263509; PubMed=10329013;  
 RA Johansson M.;  
 RT "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA  
 cloning of two novel poly(ADP-ribose) polymerase homologues.";  
 RL Genomics 57:442-445(1999).  
 RN [3]  
 RP SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=99268466; PubMed=10338144;  
 RA Berghammer H., Eder M., Marksteiner R., Auer B.;  
 RT "ADPRT-2: a novel mammalian polymerizing(ADP-ribose)transferase gene  
 related to truncated pADPRT homologues in plants and Caenorhabditis  
 elegans.";  
 RL FEBS Lett. 449:259-263(1999).  
 RN [4]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RC TISSUE=Placenta;  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,  
 Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,  
 Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K.,  
 Masuho Y., Kanehori K.;  
 RA "NED0 human cDNA sequencing project.";  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 CC -!- FUNCTION: HAS DNA-DEPENDENT POLY(ADP-RIBOSE) POLYMERASE ACTIVITY.  
 CC SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (BY  
 similarity).  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose]1(N)-acceptor -  
 nicotinamide + [ADP-D-ribose]1(N+1)-acceptor.  
 CC -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).  
 CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be  
 produced by alternative splicing.  
 CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN  
 THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO  
 DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;  
 LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND  
 THYMUS.  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AJ236912; CAB65088.1; -  
 DR EMBL; AF085734; AAD29857.1; ALT\_INIT.  
 DR EMBL; AJ236876; CAB41505.2; ALT\_INIT.  
 DR EMBL; AK001980; BAA92017.1; ALT\_TERM.  
 DR HSP; F26446; I426.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 KW ADP-ribosylation; Alternative splicing.  
 FT DOMAIN 1 88 POTENTIAL.  
 FT DOMAIN 86 583 NAD-BINDING (BY SIMILARITY).  
 FT DOMAIN 4 7 NUCLEAR LOCALIZATION SIGNAL 1ST PART  
 FT (POTENTIAL).  
 FT DOMAIN 35 40 NUCLEAR LOCALIZATION SIGNAL 2ND PART  
 FT (POTENTIAL).  
 FT VARSPLIC 68 80 MISSING (IN ISOFORM 2).  
 FT CONFLICT 447 447 P -> H (IN REF. 2).  
 FT CONFLICT 481 481 N -> H (IN REF. 4).  
 SQ SEQUENCE 583 AA; 66205 MW; 5B7AE8AE531836AF CRC64;  
 Query Match 25.1%; Score 699; DB 1; Length 583;  
 Best Local Similarity 33.8%; Pred. No. 8.1e-44;

Matches 185; Conservative 97; Mismatches 207; Indels 58; Gaps 18;  
 QY 13 GSKKQRCQTEED-----SFRSTAEALRA-----APADNRVIRVDPSPCFSPNRPG- 57  
 DB 54 GSKANKRDYEDQDGPGRSWAKRVSSEVKALLKGPAP-----VDPECTAKVGAH 106  
 QY 58 --IQVHEDYDCTLNQTNIGNNNKFFIILLEGSR--FFCWNRMGRVGEVGGSKNHFT 113  
 DB 107 VYCEGNDVYDMLNQTNLQFNKYYLIQLLEDDAQRNFVMMRWGRVGMGQHS- -VA 164  
 QY 114 C---LEDAKKDKPKFWKTKNKEEDRFVAPQPNKYTLIEVOGEAESQAVVYDGGPV 170  
 DB 165 CSGNLNKAKEIFQKKFLDKTKNNWEDREKFEKVPQKDYMLQMDYATNTQDEEETKKEESL 224  
 QY 171 RTVVKPCS--LDPATONLITNIFSKEMFKNMTIMNLNLDVKKMPLGLTKLQOIARGFEALEA 229  
 DB 225 KSLPESQDLRLVQELIKLVCNVQAMEEMEMKYNTKRAPLGLKUTVAQIRAGYOSLRK 284  
 QY 230 LEEAMKNPTGD--GQSELELSSCFYTVPHNFGSRSPPIINSPDVLOAKKMDLLVLADI 288  
 DB 285 IEDCIR--AGQHGRLMEACNEFYTRIPHDGLRTPLRTOKELSEKIQLEALGDIEI 342  
 QY 289 AQTAAQPEEEBEKVEEVPHPDLDRDYQLLRCOLQLLDGSESEYKAIQTYLKQT---GNSY 345  
 DB 343 AIKL-----VKTELQSPHPDLQHYRNLCALRLDHSYEFKVISQYLSQSTHAPTHSD 396  
 QY 346 RCPNLRHVKNVREGEDRFOAHSKLGNRRLWHGNTNVAVALTSGLRIMPH----SG 401  
 DB 397 YTMTLDDLFEVEKDGKEAFR--EDLHNRMLLWHGSRMNNVGLSHGRIAPPEAYTG 454  
 QY 402 GRVGKGIYFASNSKSAGYVYTTMHCGHGVGMFLGEVALGKHEHITIDDPKLSPPPGF 461  
 DB 455 YMGKGIYFADMSKSNKYCFASRL--KNTGLLLSEVALGQCNELEAPRAEGLQCK 512  
 QY 462 DSVIARGQTEPPADQIELELDGQPVVQGPVQCPQSPKSSSF--SQSEYLIYKESQCR 519  
 DB 513 HSTKGLGKMPSSAFV--TLNGS--TVPLGPASDTGILNPDGYTLNLYNEYIVNPNQVR 568  
 QY 520 LRVILLEI 526  
 DB 569 MRYLLUKV 575  
 RESULT 7  
 PPOL\_RAT STANDARD; PRT; 1013 AA.  
 AC P27008; O35937;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
 DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
 GN ADPRT.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;  
 RX MEDLINE=98046546; PubMed=9385436;  
 RA Beneke S., Meyer R., Buerkle A.;  
 RT "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly  
 (ADP-ribose) polymerase.";  
 RL Biochem. Mol. Biol. Int. 43:755-761(1997).  
 RN [2]  
 RP REVISION TO 811.  
 RA Beneke S., Meyer R., Buerkle A.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE OF 1-11 FROM N.A.  
 RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;  
 RX MEDLINE=92290013; PubMed=1601134;





```
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95269779; PubMed=7750552;
RA Lepiniec L., Bablychuk E., Kushnir S., van Montagu M., Inze D.;
RT "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT poly(ADP-ribose) polymerase.";
RL FEBS Lett. 364:103-108(1995).
CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC -1- SUBCELLULAR LOCATION: Nuclear
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Z48243; CAA88288.1; -
DR HSSP: P26446; 1A26.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR InterPro: IPR003034; SAP.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP; 1.
DR Pfam: PF02037; SAP; 2.
DR SMART; SM00513; SAP; 2.
DR Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation.
FT DNA_BIND 1 140 POTENTIAL.
FT DOMAIN 140 637 NAD-BINDING (BY SIMILARITY).
FT DOMAIN 41 62 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 637 AA; 72175 MW; 527A8F464605D127 CRC64;

Query Match 25.0%; Score 697.5; DB 1; Length 637;
Best Local Similarity 34.9%; Pred. No. 1.2e-43;
Matches 189; Conservative 91; Mismatches 219; Indels 43; Gaps 17;

QY 6 KASVQTEGSKKQKQGTEDSFSS--TAEALRAAPADNRVIRVDPSCPFNRNPGIQVHED 63
Db | : : : | | | | | | | | | | | | | | | | | | | | | | : | : |
112 KSSNGTDEAEDDNGGFEKKEKIVTATKGAALVDQWI---PDEIKSQYHVLQRGDD 167
QY 64 -YDCLNQTNIGNNNKFFVLIQLLEGSR--FPCNWRGVRGVEGSKMN-HFTCLEDAK 119
Db || ||||| : ||||| : || : | : ||||| ||||| : | : | : | : |
168 VYDAILNQTNVRDNNKFFVLQVLESDSKKTVMYTRGVRGKQSKLDGPDYDSDRAI 227
QY 120 KDFKKFKFEKTKNKEERDRFVAQPNKYTLIEVOGEAEQAEVKKVDSGPVRTVVKP--C 177
Db | : | : | | | | | | | | | | | | | | | | | | | | | | : | : |
228 EIFTNKNQTKNYSNDRKEFTPHPKSYTWLEMDYKKEENDSPVNDIFSSEVKAPGES 287
QY 178 SLDPAQNTIITNFSKEMFKNAWTLMLNDVKKMPGLKLTQQIARGFEALEAEAMKPN 237
Db || : : : | : | : | : | : | : | : | : | : | : | : | : | : |
288 KLDTRVAKFTSLICNYSVMAQHMGEIGYNANKLPLGKISKSTISKGYEVLRKRISEYDR- 346
QY 238 TGDQSLSELSLSCFFYIVPHNFGSRPPP--INSPDVLAQKMDMLLVLADELA-OTLQA 294
Db | ||||| ||||| : | : | : | : | : | : | : | : | : | : | : |
347 -YDRTELELSGEFYVPIHDFGKMSQFVIDTPQKLQKQIEMVEALGEIELATKLLSV 405
QY 295 APGEEEKVEEVPHPLDYDQQLLRQQLQLDLSGESVYKAIQTYLQKT---GNSYRCPNLR 351
Db || : : : | : | : | : | : | : | : | : | : | : | : | : | : |
406 DFLQDD-----PLYHYQQLNCLTPVGNDSSEFSSVWYNTMHTAKTHSGYVTEA 458
QY 352 HWKVNREGEDGRFQAHSKLGNRRLLHGHGTNVAVVAAILTSLGRIMPH-----SGGRVGKG 407
```

```
Db 459 QLFRAVRAVEADRFQFSSSKNRMLLWHGSRLLTNWAGILSQGLRIAPPEAVPTGYMFGKG 518
QY 408 IYFASSENSKAGYVVTMHC--CGHQVGVMFLGEVALGKEHHITIDDPSLKSPPGFDSV 464
Db 519 VYFADMFSSKANT-----CYANTGANDGVLLICEVALGDMNELLSDYNADNLPQKLS 573
QY 465 IARGQTEPPAPADIELELDGQPVVVPQVQCPSPKSSFSQSEYLYIKESQCRLYLL 524
Db 574 KGVGKTAPNPEAQTLLE-DG--VVVPLGRKPE-RSCSKGMLLYNEYIVNVEIKMRYVI 629
QY 525 EI 526
Db 630 QV 631

RESULT 9
PPOL_CHICK STANDARD; PRT; 1011 AA.
AC P26446;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE ADP-ribosyltransferase-1 (Poly[ADP-ribose] synthetase-1).
GN ADPRT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Oviduct;
RX MEDLINE=91340148; PubMed=1840535;
RA Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;
RT "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid
RT sequence and comparison with mammalian enzyme sequences.";
RL Gene 102:157-164(1991).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.
RX MEDLINE=96353841; PubMed=8755499;
RA Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;
RT "Structure of the catalytic fragment of poly(AD-ribose) polymerase
RT from chicken.";
RL Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO
RP 895.
RX MEDLINE=98191351; PubMed=95217110;
RA Ruf A., de Murcia G.M., Schulz G.E.;
RT "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived
RT from crystal structures and homology modeling.";
RL Biochemistry 37:3893-3900(1998).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.
RX MEDLINE=98239716; PubMed=9571033;
RA Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;
RT "The mechanism of the elongation and branching reaction of poly(ADP-
RT ribose) polymerase as derived from crystal structures and
RT mutagenesis.";
RL J. Mol. Biol. 278:57-65(1998).
CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -1- SUBUNIT: HOMODIMER (potential).
CC -1- SUBCELLULAR LOCATION: Nuclear.
```

CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X52690; CAA36917.1; -  
DR PIR; JH0581; JH0581.  
DR PDB; 2PAW; 27-MAY-98.  
DR PDB; 1PAX; 15-MAY-97.  
DR PDB; 2PAX; 27-MAY-98.  
DR PDB; 3PAX; 27-MAY-98.  
DR PDB; 4PAX; 27-MAY-98.  
DR PDB; IA26; 27-MAY-98.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP.reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP.reg; 1.  
DR Pfam; PF00645; zf-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS00172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
DR PROSITE; PS00364; PARP\_ZN\_FINGER\_2; 2.  
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.  
FT DNA\_BIND 1 370  
FT DOMAIN 371 522 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 382 458 BRCT.  
FT DOMAIN 523 1011 NAD-BINDING.  
FT ZN\_FING 21 56 PARP-TYPE.  
FT ZN\_FING 125 162 PARP-TYPE.  
FT ZN\_FING 207 209 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT MOD\_RES 403 403 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 404 404 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 410 410 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 411 411 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 432 432 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 434 434 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 441 441 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 442 442 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 453 453 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 454 454 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 468 468 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 481 481 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 485 485 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 488 488 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 509 509 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 510 510 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 517 517 ADP-RIBOSYL[N] (POTENTIAL).  
FT CONFLICT 895 895 A -> R (IN REF. 1).  
SQ SEQUENCE 1011 AA; 113520 MW; 261AED938139144 CRC64;

Query Match 24.9%; Score 694; DB 1; Length 1011;  
Best Local Similarity 34.18; Pred. No. 4e-43;  
Matches 185; Conservative 204; Indels 56; Gaps 20;  
Qy 5 RKASVQTEGSKKOROG-TEEEDSFRTAEALRAAPADNRVIRVDPSCPSRNPGIQVHED 63

Db 497 KPANMSAGKVEQPSKSEKKMKLTVKGGAAVDPDSSL--EDSAHFVEKGGKI----- 549  
Qy 64 YDCTLNQTNIGNNNKFFYIIQLLEG--SRFCWNRGRVGEV-QQSKMNFCTLEDKAK 120  
Db 550 FSATILGLVDIVKGTNSYKQLQLEDDEDRSYWVFRSGRVGTIVGSKNLEQMPKSDAVE 609  
Qy 121 DPKKFEWTKTKWEERDREVAQPNKYTLIEVOGAEQEAQVVK--VDSGPVRTVVKPCS 178  
Db 610 HFLNIEKRTNSWHSKN-FTYKPKFYPLEID-YGDEEAVRKLTVSAGTKSKLAKP-- 665  
Qy 179 LDPATQNLITIFSKEMFKNAMTLMNLDVKKMPLGKLTQQIARGFEALEALEAMKNPT 238  
Db 666 ---IQDLIKMIFDVESMKKAWFEIDLQKMLGKLSKROIQSAISILNEVQQAVSDGG 721  
Qy 239 GGGGSELSGCFYTVIHNFGSRPPPINSPDVLOAKDKMLLVLADELIAQTLQAQGE 298  
Db 722 SESQIL-DLSNRFYTLIPHDFGKMKPPLLSNLEYIQAKVQMLDNLDDIEVAYSLLRGGNE 780  
Qy 299 EEEKVEEVPHPDLDRDYQLLRQQLQLLSDGSESEYKAIQTYLKO-----TGNVSRCPNLRHVW 354  
Db 781 DGDK-----DFIDINYEKLRTDIKVVDKDSSEAKIIKQYVKNTHAATHAYDL-KVVEIF 834  
Qy 355 KVNREGEDRFQAHSKLGNRRLLNMGHTNVAVVAAIITSLGRIMPH-----SGRVRGKGIYF 410  
Db 835 RIEREGESQYKPKQLHNRQLLHGSRTTNFAGILSQGLRIAPPEAPVTGYMFKGIYF 894  
Qy 411 ASENSKSAGYVTTMCGGHQGVYMFELGVALGKEH-----HITIDDPKSLKSPGPFDSV 464  
Db 895 ADMVSKSNKYCHTSQ--ADPIGLILLGEVALGNMYELKNASHIT-----KLPGKHSV 945  
Qy 465 IARGOTEPDPAQDIELDGQPVVVPQGPVQPCPSKSSFSQSEYLYIKESQCLRYLL 524  
Db 946 KGLGKTAEDPT--ATTTLDG--VEVPLNGIS-TGINTCLLYNEYIYDVAQVNLKYL 1000  
Qy 525 EI 526  
Db 1001 KL 1002  
RESULT 10  
PPOL\_MOUSE  
ID PPOL\_MOUSE STANDARD; PRT; 1012 AA.  
AC P11103; Q9JLX4; Q9QVQ3;  
DT 01-JUL-1989 (Rel. 11, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly (ADP-ribose) polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
DE ADP-ribosyltransferase-1) (Poly(ADP-ribose) synthetase-1) (msPARP).  
GN ADPRT OR ADPRT1 OR ADPRP.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 1).  
RC STRAIN-BXSB;  
RX MEDLINE=69263780; PubMed=2498841;  
RA Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;  
RT "Sequence and organization of the mouse poly (ADP-ribose) polymerase  
RT gene.";  
RL Nucleic Acids Res. 17:3387-3401(1989).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.  
RC STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;  
RX MEDLINE=20270268; PubMed=10809783;  
RA Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;  
RT "Characterization of sPARP-1. An alternative product of PARP-1 gene  
RT with poly(ADP-ribose) polymerase activity independent of DNA strand  
RT breaks.";  
RL J. Biol. Chem. 275:15504-15511(2000).  
RN [3]  
RP KNOCK-OUT.









QY 478 IELELQGPVYVQPPVQCPSPKSSFSQSESYLIYKESQCRRLYLEI 526  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 944 Y-IRSDG--VEIPYGETITDEHLK--SSLLYNEIYVDAQVNIQYLF 988

RESULT 14  
YQ4\_CAEEL STANDARD; PRT; 538 AA.  
AC Q09525;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-FEB-1996 (Rel. 33, Last annotation update)  
DE Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.  
GN E02H1.4.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditoidea;  
OC Rhabditidae; Peloderinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Smith A.;  
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
CC -!- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-  
CC RIBOSYLTRANSFERASE (EC 2.4.2.30).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; 947075; CA887379.1; -.  
DR HSP; P26446; I426.  
DR WormPep; E02H1.4; CE01539.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 538 AA; 61268 MW; 3144E25465FC7341 CRC64;

Query Match 13.9%; Score 389; DB 1; Length 538;  
Best Local Similarity 25.9%; Pred. No. 4.le-21;  
Matches 148; Conservative 86; Mismatches 181; Indels 156; Gaps 24;  
QY 57 GIQVHEDYDCTLNQTNIGNNNKFIQLLEGRSFFCNWRGVRGCVGOSKMNHFTCLE 116  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 11 GYKVH-----LCKTNIAQNNKFFYDMLDEGGDFIVKLNIGRIGYRGVTQLKDFDDL 64  
QY 117 DAKDKPKKFWKTKNKKWEERDFVAQPNKYTLIEVQGEASQEAIV-----KVDS 167  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 65 RAKKFPESFYKTHLHWEERD-EPVPNKYAVVELATNATQTEKEVKEEPEPEKVD 123  
QY 168 GPV-----RTVVKPCSLDPATQNLTNIEFSKEMFNATFL-----MNLVKKM----- 210  
Db 124 KTRGRKKRGIVKEKEIKKEEPEVEEV--NEKLKELMKICDIEDVHLGLLQKLFNEAF 181  
QY 211 --PLGLTKQQTARGAEALAEAMKNT-----GDGQSLSELS 249  
Db 182 GRPIDLSLAQITTYEILSKTEESIGGKSARRSTRGRPRVADRVLVAKSDGPSLHDINK 241  
QY 250 CFYTVIPHNFRSRPPINSPVLAQKDMLLV-----ADIELAQLAAPGEEKEKVE 304  
Db 242 -YVSLPHSGFCVCPKIDSHAKIAERELLDALKGSIASLELDLKTASKD----- 295  
QY 305 EYVPLDRDYQLRQQLDLSGSEYKAIQIYLYKOTGNSYRCPNLR-----HWKVN----- 357  
Db 296 ----IYQRLYERLPCHEPVS-----ETAGKIGDCLAMRGPTHYCKLSLIDA 339

QY 358 -----REGEDGRFOAHSLKG-----NRRLWHGTNVAVVAALITS 392  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 340 FELKDPNEIPTAPVEVQVPPKRGKSTKTAAPTVPPTTKRLWHGRTVNVFSILMN 399  
QY 393 GLR--IMPHSGRGVKGIIYFASENSKSAGYVVTMHC--GGHQVGYMFLGEVALGK----- 443  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 400 GLQFPVGDRCGLMFGVGFANVPTKSANYC-----CPEASKRV-FMLCCEVETANPLVLY 454  
QY 444 EHHITIDPSLSKPPGFSVSTARGOTEPDPAQDIELELDQPVVVGPPVQCPSPKFS- 502  
: | | | | | : | | | | | : | | | | | : | | | | | :  
Db 455 ESEIDADEKMEKAKK--TSVYAAGKHTPRDT-----VEING-----IPAFKSN 495  
QY 503 -----SSFSQSESYLIYKESQCRRLYLEI 526  
Db 496 LETIEETRLLYDEYVNFKNKEHKIKYVVEV 526

RESULT 15  
PPOV\_HUMAN STANDARD; PRT; 1724 AA.  
AC Q9UKK3: O75903; Q9HLM6; Q14682;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault  
DE protein) (PARP-related/alphai-related H5/proline-rich) (PH5P).  
GN ADPRTL1 OR PARPL OR KIA0177.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.  
RX MEDLINE=99408776; PubMed=10477748;  
RA Kichhoefer V.A., Siva A.C., KederSha N.L., Inman E.M., Ruland C.,  
RA Streuli M., Rome L.H.;  
RT "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)  
RT polymerase.";  
RL J. Cell Biol. 146:917-928(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Thymus;  
RX MEDLINE=20112770; PubMed=10644454;  
RA Still I.H., Vince P., Cowell J.K.;  
RT "Identification of a novel gene (ADPRTL1) encoding a potential  
RT poly(ADP-ribose) transferase protein.";  
RL Genomics 62:533-536(1999).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Tromans A.;  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE OF 94-1724 FROM N.A.  
RC TISSUE=Bone marrow;  
RX MEDLINE=96281124; PubMed=8724849;  
RA Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;  
RT "Prediction of the coding sequences of unidentified human genes. V.  
RT The coding sequences of 40 new genes (K1AA0161-K1AA0200) deduced by  
RT analysis of cDNA clones from human cell line KG-1.";  
RL DNA Res. 3:17-24(1996).  
RN [5]  
RP DISCUSSION OF SEQUENCE.  
RX MEDLINE=99198702; PubMed=10100603;  
RA Jean L., Rislér J.-L., Nagase T., Coulouarn C., Nomura N.,  
RA Salier J.-P.;  
RT "The nuclear protein PH5P of the inter-alpha-inhibitor superfamily: a  
RT missing link between poly(ADP-ribose) polymerase and the  
RT inter-alpha-inhibitor family and a novel actor of DNA repair?";  
RL FEBS Lett. 446:6-8(1999).  
CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribose}(N)-acceptor =  
CC nicotinamide + {ADP-D-ribose}(N+1)-acceptor.  
CC -!- SUBUNIT: BINDS TO THE MAJOR VAULT PROTEIN (MPV). THE VAULT, A  
CC LARGE RIBONUCLEOPROTEIN COMPLEX, CONTAINS THE 100-KDA MPV AND 2





---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:01:37 ; Search time 117.38 Seconds  
(without alignments)  
778.168 Million cell updates/sec

Title: US-09-701-586b-10  
Perfect score: 2789  
Sequence: 1 MAPKKASVQTEGSKRQKG.....EYLIYKESQCLRLYLLEIHL 528

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 17299429 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2785	99.9	528	11 Q91YR6	Q91YR6 mus musculus
2	2253.5	80.8	533	4 Q96CG2	Q96CG2 homo sapien
3	823	29.5	612	5 Q9TX06	Q9TX06 dictyosteli
4	719.5	25.8	653	10 Q50017	O50017 zea mays (m
5	705	25.3	1014	11 Q921K2	Q921K2 mus musculus
6	691	24.8	635	10 Q81294	Q81294 arabidopsis
7	689	24.7	607	13 Q9P582	Q9P582 gallus gall
8	681	24.4	607	13 Q9P581	Q9P581 xenopus. na
9	624	22.4	983	10 Q9ZF54	Q9ZF54 arabidopsis
10	624	22.4	1009	10 Q9S2J4	Q9S2J4 arabidopsis
11	612.5	22.0	593	5 Q9TX05	Q9TX05 drosophila
12	584	20.9	969	10 Q24570	Q24570 zea mays (m
13	584	20.9	980	10 Q9ZSV1	Q9ZSV1 zea mays (m
14	576	20.7	945	5 Q9N4H4	Q9N4H4 caenorhabdi
15	481	17.2	727	5 Q9XUA5	Q9XUA5 caenorhabdi
16	303	10.9	2276	5 Q9TXQ1	Q9TXQ1 caenorhabdi

17	280	10.0	815	10 Q9SWB4	Q9SWB4 glycine max
18	265	9.5	815	10 Q9FK91	Q9FK91 arabidopsis
19	148	5.3	1181	5 Q9X237	Q9X237 drosophila
20	148	5.3	1181	5 Q9VBP3	Q9VBP3 drosophila
21	143	5.1	1327	4 Q95271	Q95271 homo sapien
22	131.5	4.7	363	4 Q9H8R9	Q9H8R9 homo sapien
23	121.5	4.4	924	4 Q969W4	Q969W4 homo sapien
24	121.5	4.4	1223	4 Q9UFT5	Q9UFT5 homo sapien
25	121	4.3	1092	2 Q50236	Q50236 zymomonas m
26	120.5	4.3	181	12 Q55721	Q55721 chilo iride
27	120.5	4.3	1166	4 Q9H2K2	Q9H2K2 homo sapien
28	120.5	4.3	1265	4 Q9HAS4	Q9HAS4 homo sapien
29	120	4.3	5198	5 Q76518	Q76518 caenorhabdi
30	118	4.2	935	10 Q9CA26	Q9CA26 arabidopsis
31	117	4.2	954	10 Q93826	Q93826 arabidopsis
32	115.5	4.1	451	5 Q9V635	Q9V635 drosophila
33	112	4.0	1203	16 Q9CJ19	Q9CJ19 lactococcus
34	111.5	4.0	261	4 Q9H8F2	Q9H8F2 homo sapien
35	111.5	4.0	359	4 Q9Y4P7	Q9Y4P7 homo sapien
36	110.5	4.0	1327	11 Q61595	Q61595 mus musculu
37	108.5	3.9	1849	2 Q9S4K2	Q9S4K2 lactobacill
38	108	3.9	1412	4 Q96RT1	Q96RT1 homo sapien
39	107.5	3.9	842	4 Q9HCR1	Q9HCR1 homo sapien
40	107	3.8	1330	6 Q97961	Q97961 vulpes vulp
41	107	3.8	1763	11 Q9JKK5	Q9JKK5 mus musculu
42	106	3.8	1065	2 Q9AHK7	Q9AHK7 borrelia bu
43	105	3.8	849	2 Q9AHL2	Q9AHL2 borrelia bu
44	105	3.8	1065	2 Q9AHR8	Q9AHR8 borrelia bu
45	104.5	3.7	1342	10 Q9FKN5	Q9fkn5 arabidopsis

#### ALIGNMENTS

RESULT 1

Q91YR6 ID Q91YR6 PRELIMINARY; PRT; 528 AA.  
AC Q91YR6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 59.4 KDA PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_faxid=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014870; AAL14870.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;

Query Match 99.9%; Score 2785; DB 11; Length 528;  
Best Local Similarity 99.8%; Pred. No. 6e-223;  
Matches 527; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MAPKKASVQTEGSKRQKGTEEDSFRSTAEALRAAPADNRVIRVDPSCFSPSRNPQIQV 60

|||||

Db 1 MAPKKASVQTEGSKRQKGTEEDSFRSTAEALRAAPADNRVIRVDPSCFSPSRNPQIQV 60

|||||

Qy 61 HEDYDCTLNOTNIGNNNKFFIIQLLEGRSFFCWNRRGRVGEVQSKMNFTCLEDAKK 120

|||||

Db 61 HEDYDCTLNOTNIGNNNKFFIIQLLEGRSFFCWNRRGRVGEVQSKMNFTCLEDAKK 120

|||||

Qy 121 DFKKKFEKTKNKEERDFVAQPNKYTLIEVQGAESQEAQVVKVDSGPVTVVKPCSLD 180

|||||

Db 121 DFKKKFEKTKNKEERDFVAQPNKYTLIEVQGAESQEAQVVKVDSGPVTVVKPCSLD 180

|||||

Qy 181 PATONLITNIFSKEMFKNMTLMNLDVKKMPLGLTKQIARGFPALEAEAKNPTGD 240

|||||

```
Db 181 PATQNLITNIFSKEMFNKMTLMNLVDVKKMPLGKLTQQTARGFALAELEAMKNPTGD 240
QY 241 GQSLLEELSSCFYTVIPHNFRGRPPPIPSDVLQAKKMDLLVLADIATLQAAPGEEE 300
Db 241 GQSLLEELSSCFYTVIPHNFRGRPPPIPSDVLQAKKMDLLVLADIATLQAAPGEEE 300
QY 301 EKVEEVPHPDLDYQQLRQQLLDGSESEYKAIQTYLKTGNSYRCPNLRHVKVNVREG 360
Db 301 EKVEEVPHPDLDYQQLRQQLLDGSESEYKAIQTYLKTGNSYRCPNLRHVKVNVREG 360
QY 361 EGRFQAHSLGNRLRLWHTGNVAVAAITSLGRIMPHSGSGRGVKGIIYFASENKSAGY 420
Db 361 EGRFQAHSLGNRLRLWHTGNVAVAAITSLGRIMPHSGSGRGVKGIIYFASENKSAGY 420
QY 421 VTMHCGGHGVYMFGLGEALGKEHHITIDDPKSLKPPPGFSDVIARGQTEPPDPAQDI 480
Db 421 VTMHCGGHGVYMFGLGEALGKEHHITIDDPKSLKPPPGFSDVIARGQTEPPDPAQDI 480
QY 481 ELDGQPVVPPQGPVQCPSPKSSFSQSEYLIYKESQCRRLYLLEIHL 528
Db 481 ELDGQPVVPPQGPVQCPSPKSSFSQSEYLIYKESQCRRLYLLEIHL 528

RESULT 2
Q96CG2
ID Q96CG2 PRELIMINARY; PRT; 533 AA.
AC Q96CG2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 60.1 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014260; AH14260.1; -.
KW Hypothetical protein.
SQ SEQUENCE 533 AA; 60089 MW; 6296A0E439CC7767 CRC64;
```

```
Query Match 80.8%; Score 2253.5; DB 4; Length 533;
Best Local Similarity 80.9%; Pred. No. 9.4e-179;
Matches 432; Conservative 37; Mismatches 58; Indels 7; Gaps 4;

QY 1 MAPKKASVQTEG--SKKQKQGTTEEDSFRSTAEALRAAPADNRVIRVDPSCPSRNPGI 58
Db 1 MAPKPKPVQTEGPEKKKGQAGREDDPFSTAEALKAIPAERIRVDPCTPLSSNPGT 60

QY 59 OVHEDYDCTLNQINIGNNKFIYIQLLGRSRRF-CWNRWGRVGEVQSGKMHTFCTLED 117
Db 61 QVYEDYACTLNQINIGNNKFIYIQLLQDSNRRFTCWNRWGRVGEVQSGKINHTFLED 120

QY 118 AKKDFKKKFEKTKNWEERDRFVAQPNKYTLLEVGAEASQEAQVVKVDSGPVTV--V 174
Db 121 AKKDFKKKFEKTKNWEERDRFVAQPNKYTLLEVGAEASQEAQVVKVDSGPVTVTKRV 180

QY 175 KPCSLDPATONLITNIFSKEMFNKMTLMNLVDVKKMPLGKLTQQTARGFALAELEAM 234
Db 181 QPCSLDPATOKLTNIFSKEMFNKMTLMNLVDVKKMPLGKLTQQTARGFALAELEAM 240

QY 235 KNPFGQGSLEELSSCFYTVIPHNFRGRPPPIPSDVLQAKKMDLLVLADIATLQA 294
Db 241 KNPFGQGSLEELSSCFYTVIPHNFRGRPPPIPSDVLQAKKMDLLVLADIATLQA 300

QY 295 APGEEEKVEEVPHPDLDYQQLRQQLLDGSESEYKAIQTYLKTGNSYRCPNLRHVK 354
Db 301 V-SEKTEVEEVPHPDLDYQQLRQQLLDGSAPEYKVIQTYLQETGSHNRCPTLQIHW 359
```

```
QY 355 KVNREGEDRFQAHSLGNRLRLWHTGNVAVAAITSLGRIMPHSGSGRGVKGIIYFASEN 414
Db 360 KVNREGEDRFQAHSLGNRLRLWHTGNVAVAAITSLGRIMPHSGSGRGVKGIIYFASEN 419
QY 415 SKAGYVVTMHCGGHGVYMFGLGEALGKEHHITIDDPKSLKPPPGFSDVIARGQTEPDP 474
Db 420 SKAGYVVTMHCGGHGVYMFGLGEALGKEHHITIDDPKSLKPPPGFSDVIARGQTEPDP 479
QY 475 AQDIELELDGQPVVPPQGPVQCPSPKSSFSQSEYLIYKESQCRRLYLLEIHL 528
Db 480 AQDIELELDGQPVVPPQGPVQCPSPKSSFSQSEYLIYKESQCRRLYLLEIHL 533

RESULT 3
Q9TX06
ID Q9TX06 PRELIMINARY; PRT; 612 AA.
AC Q9TX06;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
SQ SEQUENCE 612 AA; 69241 MW; CB340F7A88BF2364 CRC64;
```

```
Query Match 29.5%; Score 823; DB 5; Length 612;
Best Local Similarity 38.4%; Pred. No. 9.9e-60;
Matches 209; Conservative 79; Mismatches 204; Indels 52; Gaps 17;
```

```
QY 6 KASVQTEGSKKQKQGTTEEDSFRSTAEALRAAPADNRVIRVDPSCPSRNPGIQVHED-- 63
Db 96 EAEKTAASDLDDSSSESEDEKNOISVKIKGRAAN-----DPHFPSR---XHYVENGK 146

QY 64 --YDCTLNQINIGNNKFIYIQLLE--EGSRFFCWNRWGRVGEVQSGKMHTF--TCLED 117
Db 147 DVYDATLNQTEIISONNNKXYIIQLLEADGSSYSVWNRWREGLGKSSSRKDFGKGLNQ 206

QY 118 AKKDFKKKFEKTKNWEERDRFVAQPNKYTLLEVGAEASQEAQVVKVDSGPVTVV 174
Db 207 AISLFCSKFEKTKNFTDRANFKVAGKYDMLDYSTDSPKPKNGASTATTATTTTKVV 266

QY 175 ---KPCSLDPATONLITNIFSKEMFNKMTLMNLVDVKKMPLGKLTQQTARGFALAELE 231
Db 267 EHKKECSLDERVELVKLIPDVKKMERTMTEAKYDLKMPGLKSLKNQITKGLVLUKQIE 326

QY 232 EAMKNTPDGQGSLEELSSCFYTVIPHNFRGRPPPIPSDVLQAKKMDLLVLADIATLQA 291
Db 327 DVMGGS--GESLSTLSSRFYTIIPHAFGMSVPPVINTNQMLIEKMNMLQNLADIEIATN 384

QY 292 LQAAPGEEEEKVEEVPHPDLDYQQLRQQLLDGSESEYKAIQTYLKTGNSYRCPNLR 351
Db 385 IIKDSESDSNI-----LELHYAKLKTDIQPLDENSECYKNLLYKNTYQGGKPTIV 438

QY 352 HWKVKVNRGEGRFQAHSLGNRLRLWHTGNVAVAAITSLGRIMPH-----SGGRVGK 407
Db 439 NIFKIDRDEADRIYKTKHGNKLLWHSGRJNYASIIISQGLRIAPPEAPVSGYRFGK 498
```

Qy	356	VNREGEDGRFOAHKSLGNRRLLWHGTINAVVAAAILTSLGLRIMPH-----SGGRVGKGIIYFA	4111
Dq	356		
Db	478	VSRHGETERFQKFASTRNMLLWHGSRLSNWAGLTSOGLRIAPPEAAPTGYMFGKGVYFA	5377
Qy	412	SENSKSAGRYVTTHCGGHQV---CYMFLGEVALGKEHHIITDDPSLAKSPPGDSDVIARG	4688
Dq	412		
Db	538	DMFSKSANY-----CYASEACRSGLLVCEVALGDMNMELLNADYDANNLPGKLRSKGVG	5929
Qy	469	QTEPDPAQIDLELDGPVPVPPQPCPQSFSSSFSQSEYLIYIKESQCRLRYLLEIH	5277
Dq	469		
Db	593	QTAPNM---VESKVADDGVVVPLGEPKQEPS-KRGGLLYNEYIVYNVDQIRMYRVVLHVN	6477
RESULT	5		
Q921k2		PRELIMINARY;	PRT; 1014 AA.
ID	Q921K2		
AC	Q921K2;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last annotation update)		
DE	SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE) POLYMERASE).		
DE	POLYMERASE).		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
OX	NCBI_TaxID=10090;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RA	Strausberg R.;		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; BC012041; AALH2041.1; "		
KW	Transferase.		
SQ	SEQUENCE 1014 AA; 112721 MW; 5DCE68E4CB3F46EB CRC64;		
Query Match		25.3%;	Score 705; DB 11; Length 1014;
Best Local Similarity		33.6%;	Pred. No. 1.4e-49;
Matches		182; Conservative	101; Mismatches 214; Indels 44; Gaps
Qy	1	MAPKKASVOTEGSKKORQGTEEDSPFRSTAEALRAAPADNRVIRVDPSCPFRRNPQIQV	60
Dq	1		
Db	494	VAPKGSAAPS KSKS----GAVKEEGVKNSEKRKMKLT LKGAAVDPDPSGLESHAVHLEKG	5497

[illegible]



Matches	178;	Conservative	95;	Mismatches	215;	Indels	42;	Gaps
Qy	10	QTEGSKKQROGT--EEDSFSTAEALRAAPADNRVRVDPSPCFSRNPQGVHEDYDCTL 68						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	97	KSSGKVKEEGSKSEKKMKLTVKGGAAIDPDS---ELEDSCHVLETGG---KIFSATL 149						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	69	NOTWIGNNNKFYLIQLE--EGSRFCFNNRWGRGVEV--GQSKMNIFTCLEDAKKDFKKK 125						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	150	GLVDITRGTSYTKQLIEHDSDRSRYWFRSGVRGVTGCSKLEEMSKEDAIEHFLNL 209						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	126	PWEKTKNKEERDFEVAQPNKTYLIEVQ--GEAESQEAIVKDSGPPVTVVVKPCSLDPAQ 184						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	210	YODKTGNAWHS--PNFTKYPKKFPLEIDYQGEEDVVKVLSVGAGTKSKLAKP-----VQ 262						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	185	NLIITNIESKEMFNAMTLMNDLVKKMPLGKLTQQIARGFEALEALEEAMKNPTGDQSL 244						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	263	ELIKLIFDVESMKKAWVEFIDQLKMPGLKSKRQIQSAYSILSQVQAVSESLSEARLL 322						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	245	EELSCSYTVIPINFGSRPPPTNSPDVLOAKKMDLLVLADIETLAOTLOAAPEGEEBKVE 304						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	323	-DLSNQFYTLIPHDFGMMKPLPNNLEYIQAKVQMLDNLIDIEAVYSLLRGGAADGGEK-- 379						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	305	EVPHPLDRDYQLLRCOLQLDSESEYKAIQTLYK----QTGNSYRCPNLRHWKVNREG 360						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	380	---DPIDVKEKTKTDIKVYAKDSESRIFCDVYKNTHADTHNAYDLEVL--EIFIKIDREG 435						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	361	EGDRFOAHSKLGNNRLNLWHTGNVAVAAILTSLGRIMPH-----SGRGVKGITYFASENSK 416						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	436	EYQRYKFPKQLHNRQLLWHSRRTNFAGILSQGLRTAPEAPVTGYMFGKGIYFADWYSK 495						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	417	SAGYVTTVMHCGGHQGVYMFGLGEVALGKEHHITIDDPSLKSPPPGFSVIARGQTEPDPAQ 476						
		:   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	496	SANYCHAM--PGSPIGLILLGEVALGNMHELKAASQITKL--PKGKSHVKGIGRTAPDPSPA 552						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Qy	477	DIELELDGQPVVVPQGPVQCPSEKSSFSQSESYLYIKESQCKRLRYLLEI 526						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						
Db	553	TV--QLDG--VDVPLAKGTSA--NISDTSLLYNEYIYDIAQVNLKLYLKKL 597						
		: :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :   :						

```

RESULT      9
Q9ZP54      PRELIMINARY;      PRT;      983 AA.
AC      Q9ZP54;      Q9ZP54;
DT      01-MAY-1999 (TtEMBLrel. 10, Created)
DT      01-DEC-2001 (TtEMBLrel. 19, Last sequence update)
DT      01-DEC-2001 (TtEMBLrel. 19, Last annotation update)
DE      POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN      PARP-1.
OS      Arabidopsis thaliana (Mouse-ear cress).
OC      Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC      Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC      eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX      NCBI_TaxID=3702;
RN      [1]
RC      SEQUENCE FROM N.A.
RC      STRAIN=CV. LANDSBERG RECTA;
RA      Doucet-Chabeaud G., Kazmaier M.;
RL      Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AJ131705; CAA10482.1; -.
DR      HSSP; P26446; 1A26.
DK      Transferase; Glycosyltransferase; NAD.
FT      CHAIN      2      983      POLY(ADP-RIBOSE) POLYMERASE.
SQ      SEQUENCE      983 AA;      111232 MW;      468E12A8EF1B6F4F CRC64;

Query Match      22.4%;      Score 624;      DB 10;      Length 983;
Best Local Similarity      33.0%;      Pred. No. 7.2e-43;
Matches 179;      Conservative      87;      Mismatches 209;      Indels      68;      Gaps

QY      15      KKQQTGEEDSFRTAEALRAPADNRVIRVDPCFFSNPQIVH-----ED----YDC 66
      ||||      ::      |::|      |      |      |      |      |      |
Db      472      KQKKLPDFDKYKIEDTSVLTVVKVGRSAVHEAS-----GLOGHCHILEDGNSINYNT 524

```

Query Match	24.4%	Score 681;	DB 13;	Length 607;
Best Local Similarity	33.6%	Pred. No. 6.3e-48;		

```
QY 67 TLNQTIGNNNNFYIIQLLEE--GSRFCWNRWGRV--EVQSKMNHFTCLDEAKKDF 122
Db 525 TLSMSDLSTGINSYYIIQIEDKSDCYVFRKGRVGNKEIGNKVEEMS--KSDAVHEF 593
QY 123 KKKFWKTKN---KWEERDRFVAQPNKYTLIEVQBAESQAEVVKVDSGVPRTVVKPCSL 179
Db 584 KRLFEKTGNTWESWEQKTNFQKPGKFLPLDIDYGVNKQVA-----KKEPFQT---SSNL 636
QY 180 DPATQNLITNIFSKEMFKNMTLMNLDVKKMPLGKLTQKQIARGFEAL---EALSEAMK 235
Db 637 APSLIELMKMLFDVETYSAMMEFEINMSEMPGLKSKHNIQKGFALTEIQRLTESDP 696
QY 236 NPTGDCQSLEELSSCFYTVIPHNFGRSRPPPIINSPVQLAOKKMDMLVLADIOLAQTL 295
Db 697 OPTMKESLLVDASNRFTMP-----SIHPHIIREDDEFSKVKMLQALQDIETASRI--- 749
QY 296 PGEEEEKVEVPHPLDRDYQLLRQQLDGSSEYKATQTYLKQGNVSRCP----- 348
Db 750 VGFVDVSTES---LDDKYKKLHCDISPLHDSYEDYRLIEKYL---NTHAPTHTWSL 801
QY 349 NLRHVWVKNREGEDRFQAH--SKLGNRRLLWHGNTNVAVAAILTSLRI-----MPHSGGR 403
Db 802 ELSEVFALEGEFDKYAPHREKLGKMLLWHSRLTNFVGLINQGLRIAPPAPATGYM 861
QY 404 VKGIIYFASSENSAGYVTTMHCGGHQGVYMFGLGEVALGKEHHITIDDPSSLKSPPPGDFS 463
Db 862 FGKGIYFADLVSKSAQCYIT--CKKNPVGLMLLSEVALGEIHELT--KAKYMDKPPRGKHS 918
QY 464 VIARGOTEPDPAQDIELELDGQPVVPPQGPVQCPGKSSFSQSEYLYIKESQCRRLYL 523
Db 919 TRGLGKKVP---QDSEFAKRWGDTVPCGKPVSV--SKVKASELMYNEYIYVDTAQVKLQFL 974
QY 524 LEI 526
Db 975 LKV 977

RESULT 10
Q9SJW4 PRELIMINARY; PRT; 1009 AA.
ID Q9SJW4;
AC Q9SJW4;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN AT2G31320.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Unayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eissen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana."
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006593; AD20677.1; -.
DR HSSP; P26446; IA26.
DR InterPro; IPR001357; BRCT.
'.
```

```
DR InterPro; IPR001390; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; ZnF-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; ZnF-PARP; 2.
DR ProDom; PD004675; ZnF-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS50172; BRCT; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ SEQUENCE 1009 AA; 114133 MW; CDEGE41CC2A3A2DB CRC64;

Query Match 22.4%; Score 624; DB 10; Length 1009;
Best Local Similarity 33.0%; Pred. No. 7,5e-43;
Matches 179; Conservative 87; Mismatches 209; Indels 68; Gaps 21;

QY 15 KKQRCQTEEDSFRSTAEALRAAPADNRVIRVDPSPFSSRNPGIQVH-----ED----YDC 66
Db 498 KKQKRLPFQDKYKIEDTSLSIVTVKGRSAVHEAS-----GLQEHCHILEDGNSIYNT 550
QY 67 TLNQTIGNNNNFYIIQLLEE--GSRFCWNRWGRV--EVQSKMNHFTCLDEAKKDF 122
Db 551 TLSMSDLSTGINSYYIIQIEDKSDCYVFRKGRVGNKEIGNKVEEMS--KSDAVHEF 609
QY 123 KKKFEWTKN---KWEERDRFVAQPNKYTLIEVQBAESQAEVVKVDSGVPRTVVKPCSL 179
Db 610 KRLFEKTGNTWESWEQKTNFQKPGKFLPLDIDYGVNKQVA-----KKEPFQT---SSNL 662
QY 180 DPATQNLITNIFSKEMFKNMTLMNLDVKKMPLGKLTQKQIARGFEAL---EALSEAMK 235
Db 663 APSLIELMKMLFDVETYSAMMEFEINMSEMPGLKSKHNIQKGFALTEIQRLTESDP 722
QY 236 NPTGDCQSLEELSSCFYTVIPHNFGRSRPPPIINSPVQLAOKKMDMLVLADIOLAQTL 295
Db 723 OPTMKESLLVDASNRFTMP-----SIHPHIIREDDEFSKVKMLQALQDIETASRI--- 775
QY 296 PGEEEEKVEVPHPLDRDYQLLRQQLDGSSEYKATQTYLKQGNVSRCP----- 348
Db 776 VGFVDVSTES---LDDKYKKLHCDISPLHDSYEDYRLIEKYL---NTHAPTHTWSL 827
QY 349 NLRHVWVKNREGEDRFQAH--SKLGNRRLLWHGNTNVAVAAILTSLRI-----MPHSGGR 403
Db 828 ELEEVFALEGEFDKYAPHREKLGKMLLWHSRLTNFVGLINQGLRIAPPAPATGYM 887
QY 404 VKGIIYFASSENSAGYVTTMHCGGHQGVYMFGLGEVALGKEHHITIDDPSSLKSPPPGDFS 463
Db 888 FGKGIYFADLVSKSAQCYIT--CKKNPVGLMLLSEVALGEIHELT--KAKYMDKPPRGKHS 944
QY 464 VIARGOTEPDPAQDIELELDGQPVVPPQGPVQCPGKSSFSQSEYLYIKESQCRRLYL 523
Db 945 TRGLGKKVP---QDSEFAKRWGDTVPCGKPVSV--SKVKASELMYNEYIYVDTAQVKLQFL 1000
QY 524 LEI 526
Db 1001 LKV 1003

RESULT 11
Q9TX05 PRELIMINARY; PRT; 593 AA.
ID Q9TX05;
AC Q9TX05;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NAD+ PROTEIN (ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS Drosophila sp. (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7242;
RN [1]
```



```
RP SEQUENCE FROM N.A.
RX MEDLINE-96007847; PubMed-7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RT Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+; protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostellium discoideum and
RL inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSSP; P26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR PROSITE; PS0172; BRCT; 1.
SQ SEQUENCE 593 AA; 68018 MW; D9BA37E38B8E7CCD CRC64;

Query Match 22.0%; Score 612.5; DB 5; Length 593;
Best Local Similarity 31.4%; Pred. No. 3e-42;
Matches 166; Conservative 102; Mismatches 222; Indels 39; Gaps 17;

Qy 16 KQGTEDSFRSTAELRAAPADNRVIRVDPSCPSRNPQIQ-----VHED-----YDCT 67
Db 80 KSRIPKETKSLNSNSIYTKSPV-SRTFKYKDGGLAVDPDPSGLEIDIAHYVYDSNNKYSV 138
Qy 68 LNOTNIGNNNKXYLIOL--EESRRFCNWRGVG--EYGQSKMNHFTCLEDAKKDFKK 124
Db 139 LGLTDIQRNKSXYKVLKADKKKEKWFIRSGRIGNTGNISKLEEFDTSESAKRNFK 198
Qy 125 KFEKTKNKEERDRFAQPNKYTLIEVQGEAESQEAQVVDGSPVIRVVKPCSLDPATQ 184
Db 199 IYADKTGNEYQDRNFVKTGMYPIELQ--YDDQKLVKHSHFTS-----KLEISVQ 251
Qy 185 NLITNIFSKEMFNAMTLMNDLVKMPGLKLTQKQIARGFALPALEAFAMKNPTGDGQSL 244
Db 252 NLILIEDIDSMNKTLEFHDMDKMPGLKLSAQHIOQSAVVRKEIYNVLECGSNTAK-L 310
Qy 245 EELSCFVTVPHNFGSRPPINSVDVLAQKDMLLVLADELIAQTLQAAPGEEEEKVE 304
Db 311 IDATNRFYTLPHNFGVOLPTLIETHQOIEDLRQMLDSLAIEVAYST-----IKSEDVS 365
Qy 305 EVPHPLRDYQLLRQQLDGSSEYKAQIOTYIKQT--GNSYRCPNLR--HWKVNREGE 361
Db 366 DACNPLDNHYAQIKTQVLDKNSSESLSQYVKNTHASTHKSVDLKIYDVFAVSRQGE 425
Qy 362 GDRFOAHSKLGNRLLWHGTNVAVVAAILTSGRLI----MPHSGRGVKGKGIYFASENSKS 417
Db 426 ARREKPFKKLHNRLLWHGSRLTNFGVILSHGLRIAPPEAPTGYMFGKGIYFADMYSKS 485
Qy 418 AGYVVTMHCGGHOVGMFLGEVALGKEHHITIDPDLKSPPGFDSVIARGOTEPDPAQD 477
Db 486 ANYCCTSQ--QNSTGLMLLSEVALGDMMWECT-SAKYINKLSNNKHSFCGRGTMPDPTKS 542
Qy 478 IELELDGPPVVPVPPQCPSPKSSFSQSEYLIYKESQCRRLYLEI 526
Db 543 Y-IRSDG--VEIPYGETITDHLK-SLLIYNEYIYDVAQVNIQYLFEM 587

RESULT 12
ID 024570 PRELIMINARY; PRT; 969 AA.
AC 024570;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLY(ADP-RIBOSE) POLYMERASE.
GN PARP.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
```

```
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Babychuk E., Cottrell P., Storozhenko S., Fuangthong M.,
RA O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT "Higher plants possess two poly(ADP-ribose) polymerases.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222589; CAA10889.1; -.
DR HSSP; P26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR003034; SAP.
DR InterPro; IPR001510; ZnF-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; ZnF-PARP; 2.
DR ProDom; PD004675; ZnF-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS0172; BRCT; 1.
DR PROSITE; PS00064; PARP_ZN_FINGER_2; 2.
SQ SEQUENCE 969 AA; 109128 MW; EB23AC62EEC14009 CRC64;

Query Match 20.9%; Score 584; DB 10; Length 969;
Best Local Similarity 30.9%; Pred. No. 1.5e-39;
Matches 164; Conservative 94; Mismatches 201; Indels 72; Gaps 20;

Qy 33 ALRAAPADNRVIRVDPSCPSRNPQIQ-----VHED-----YDCTLNQTNIGNNNKFIYIQ 84
Db 468 ALESSKGTSTVTKVKGSAVHSESSGLQDTHAILEDGKSIYNATLMSDLALGVSYYVLQ 527
Qy 85 LLE--EESRRFCNWRGVG--EYGQSKMNHFTCLEDAKKDFKKFKWEKTKNKEE--R 137
Db 528 IIEQDDGSECVYFKWGRVSGEIKGGQKLEEMSKTE-AIKEFKRLFLKTKGNSWEAMECK 586
Qy 138 DRFAVQPNKYTLIEVQGEAESQEAQVVDGSPVIRVVK--PCSLDPATQNLITNIFSKEM 195
Db 587 TNFRKQGRFYPLDVG-----YGVKKAPKRDISEMKSSSLAPOLLELMKMLFNVET 637
Qy 196 FKNAMTLMNDLVKMPGLKLTQKQIARGFALPALEAFAMKNPTGDGOSLE-----LSSC 250
Db 638 YRAAMEFEINSMPLKLSKENIEKGFALTEIQNLKDTADQALAVRESLIVAASNR 697
Qy 251 FYTVIPHNFGRSRPPINSVDVLAQKDMLLVLADELIAQTLQAAPGEEEEKVEVPHPL 310
Db 698 FFTLIP-----SIHPHIIIRDEDDLMKAKMLEALQDIETASKIVGFDSDSDS-----L 746
Qy 311 DRDYOLLRCQLDLSGESEYKAQIOTYIKQTGNSYRCP-----NLRHVKNVREGE 363
Db 747 DDKYMKLHCDITPLAHSDSDYKLLIEQYLLNT-----HAPTHKDWSELEEEVSLORDGELN 802
Qy 364 RFQAH-SKLGNRLLWHGTNVAVVAAILTSGRLIMPH-----SGGRVKGKGIYFASENSKSA 418
Db 803 KYSRYKNLHKNMLLWHGSRLTNFGVILSGLRIAPPEAPTGYMFGKGLYFADLVSKSA 862
Qy 419 GYVVTMHC--GGHOVGMFLGEVALGKEHHITIDPDLKSPPGFDSVIARGOTEPDPA 475
Db 863 QY-----CYDRNNPVGMLLSEVALGDMDYELK-KATSMDKPPPKGKHSKGLGTVPLES 916
Qy 476 QDIELELDGPPVVPVPPQCPSPKSSFSQSEYLIYKESQCRRLYLEI 526
Db 917 EFVKWRDD--VVVPCGKVPV-SIRSSELMYNEYIYNTSOVKMQFLKV 963

RESULT 13
Q9ZSV1
ID Q9ZSV1 PRELIMINARY; PRT; 980 AA.
AC Q9ZSV1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
```





**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:57:19 ; Search time 124.84 seconds  
(without alignments)  
469.777 Million cell updates/sec

Title: US-09-701-586B-10

Perfect score: 2789

Sequence: 1 MAPKKKASVQTEGSKKQKOG.....EYLIVKESQCLRYLLEIHL 528

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2789	100.0	528	21 AAY51178	Murine PARP1 (short)
2	2776.5	99.6	533	21 AAY51177	Murine PARP1 (long)
3	2253.5	80.8	533	21 AAY51175	Human brain PARP3
4	2253.5	80.8	540	21 AAY51176	Human uterus type
5	2241.5	80.4	533	22 AAU29021	Human PARP-3 prote
6	719.5	25.8	653	21 AAY68834	A poly(ADP-ribose)
7	707.5	25.4	1013	17 AAR98642	Poly(ADP-ribose) p
8	707	25.3	522	22 AAU29022	Mouse PARP-2 prote
9	702.5	25.2	1014	21 AAY58043	Human poly (ADP-ri
10	702.5	25.2	1014	22 AAU29019	Human PARP-1 prote
11	702.5	25.2	1014	22 AAB66296	Human tankyrase2 r

12	701.5	25.2	1014	20 AAY33699	Human poly(ADP-rib
13	700.5	25.1	521	22 AAB60693	Human poly(ADP-rib
14	700.5	25.1	570	21 AAY51174	Human brain PARP2
15	700.5	25.1	570	22 AAB11480	Human brain poly-A
16	699	25.1	534	21 AAB42909	Human ORFX ORF2673
17	699	25.1	534	22 AAU29023	Human PARP-2 prote
18	699	25.1	583	22 AAB47029	hPARP2. Homo sapi
19	698.5	25.0	1014	21 AAY49939	Human nuclear NAD+
20	698	25.0	534	22 AAU29020	Human PARP-2 prote
21	697.5	25.0	637	21 AAY68835	The poly(ADP-ribos
22	649	23.3	531	22 AAB93513	Human protein sequ
23	621.5	22.3	557	22 AAB66431	Drosophila melanog
24	609	21.8	1063	22 AAB47032	Fusion protein PAR
25	584	20.9	969	21 AAY68833	A poly(ADP-ribose)
26	584	20.9	980	21 AAY68839	A poly(ADP-ribose)
27	490.5	17.6	982	20 AAY28464	Maize poly ADP-rib
28	456.5	16.4	379	22 AAU21687	Novel human neopla
29	437.5	15.7	360	22 AAB47030	N-terminal fragmen
30	405.5	14.5	1010	21 AAY68840	Fusion protein of
31	378.5	13.6	294	22 AAU20129	Human DNA repair a
32	378.5	13.6	294	22 AAU21810	Novel human neopla
33	364.5	13.1	287	22 AAB47031	C-terminal fragmen
34	276.5	9.9	1730	22 AAU33242	Novel human secret
35	272	9.8	1724	21 AAY54373	cDNA sequence enco
36	272	9.8	1724	22 AAB51022	Human minor vault
37	267.5	9.6	1099	22 AAB66301	Human tankyrase2 e
38	246.5	8.8	227	22 AAU20130	Human DNA repair a
39	246.5	8.8	227	22 AAU21811	Novel human neopla
40	203.5	7.3	190	22 AAU21688	Novel human neopla
41	148	5.3	1181	22 AAB60894	Drosophila melanog
42	148	5.3	1181	22 AAB66297	Drosophila tankyra
43	143	5.1	1327	21 AAB27212	Human tankyrase I
44	143	5.1	1327	21 AAY44402	Human tankyrase.
45	143	5.1	1327	22 AAB66279	Human tankyrase1 S

#### ALIGNMENTS

#### RESULT 1

AAU51178  
ID AAY51178 standard; Protein; 528 AA.

XX  
AC AAY51178;

XX  
DT 31-MAR-2000 (first entry)

XX  
DE Murine PARP1 (short) homologue protein.

XX  
KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
XX  
KW ischemic tissue damage; PARP1.

XX  
OS Mus sp.

XX  
PN WO9964572-A2.

XX  
PD 16-DEC-1999.

XX  
PF 04-JUN-1999; 99WO-EP03889.

XX  
PR 05-JUN-1998; 98DE-1025213.

XX  
PR 01-MAR-1999; 99DE-1008837.

XX  
PA (BADI ) BASF AG.

XX  
PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;

XX  
DR WPI; 2000-087218/07.

XX  
DR N-PSDB; AAZ44291.

XX  
PT Novel genes and proteins, antibodies and binding partners useful in  
diagnosis and therapy of energy deficiency associated disease

conditions -  
 Claim 4; Page 71-73; 96pp; German.  
 This invention describes novel human and murine poly(ADP-ribose) polymerase (PARP) homologues, which are characterised by an amino acid sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MHX<sub>2</sub>C (I). The nucleic acid sequences, PARP homologues and antibodies are useful for analytic detection of PARP homologues and for identifying PARP effectors or binding partners, as well as for determining their effectiveness. PARP-binding partners are useful for the diagnosis or therapy of a disease condition, which is the result of a PARP protein, especially an energy deficiency, which may comprise tissue damage from cell death following necrosis or apoptosis. The disease condition may be chosen from a neurodegenerative illness, or sepsis or ischemic tissue damage, in particular neurotoxic disturbances, etc. This sequence represents the murine PARP1 protein used in the method of the invention.

Query Match 100.0%; Score 2789; DB 21; Length 528;  
 Best Local Similarity 100.0%; Pred. No. 4.3e-239;  
 Matches 528; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAPKRASVQTEGSKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIQV 60  
 DB 1 mapkrasvqtegskqrogteedsfrstaealraapadnrvirvdpscfgrnpgiqv 60  
 QY 61 HEDYDCTLNQNTGNNGNNFYIIQLLEGSFRFCWNRWGRVGEVGSKMNHFTCLEDAKK 120  
 DB 61 hedydctlnqntgnngnnfyiqlleegsfrfcwncwgrvgrvgevggskmnhftcleddakk 120  
 QY 121 DFKKKFEKTKNWEERDRVAPQNKYTLIEVQGEAESQAVVK---VDSGPVRTVVK 180  
 DB 121 dfkkkfwetkknweerdrrvapgnytlievqgeaesgeavvkvdsqdvsgpvrvtvk 180  
 QY 181 PATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQOIARGFEALAEAMKNPTGD 240  
 DB 181 patqnlitnifskemfnamtlmnlvdkmpltgkltqoiargfealealeeamknptgd 240  
 QY 241 GQSLLELSSCFYVIPHNFGRSRPPPIINSPDVLOAKKDMLLVLADIETLAQTQAAGPEEE 300  
 DB 241 gqsllelsscfyvipnfngrsrpppnsdpvlqakkdmlvlvadielaqtlqaagpeee 300  
 QY 301 EKVEEVPHPDLRDYQLLRCOLQLDSESEYKATQTYLKTGNSYRCNLRHVKNREG 360  
 DB 301 ekveevphpldrdyqlrrcqlldsgeseykaitylktgnsyrcpnlrhvknvreg 360  
 QY 361 EGRFQAHSLGNRRLLWHTNVAVAAILTSGLRIMPHSGRGVKGITYFASENSKSAGY 420  
 DB 361 egdrfqahslgnrrllwhtnvaavaailtsglrimphsgrgvgkityfasensksagy 420  
 QY 421 VTMHCGGHGVWFGEVAGLKEHHTIDDDPSLKSPPPGFDVSIARGQTEPPDAQDIEL 480  
 DB 421 vtmhcgghgvwmfgevalgkehhtiddpslskspppgfdsviargqteppdaqdiel 480  
 QY 481 ELDGQPVVPQGPVQCPSPFKSSFSQSEVLIYKESQCLRLYLLEIHL 528  
 DB 481 eldgqpvvpqgpvpvcpsfkssfsqseyllykesqclrlylleihl 528

RESULT 2  
 ID AAY51177  
 AC AAY51177  
 DT AAY51177;  
 XX 31-MAR-2000 (first entry)  
 XX Murine PARP1 (long) homologue protein.

PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis; diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness; ischemic tissue damage; PARP1.  
 Mus sp.  
 W09964572-A2.  
 16-DEC-1999.  
 04-JUN-1999; 99WO-EP03889.  
 05-JUN-1998; 98DE-1025213.  
 01-MAR-1999; 99DE-1008837.  
 (BADI ) BASF AG.  
 Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
 WPI: 2000-087218/07.  
 N-PSDB; AA244290.  
 Novel genes and proteins, antibodies and binding partners useful in diagnosis and therapy of energy deficiency associated disease conditions -  
 Claim 4; Page 67-69; 96pp; German.  
 This invention describes novel human and murine poly(ADP-ribose) polymerase (PARP) homologues, which are characterised by an amino acid sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MHX<sub>2</sub>C (I). The nucleic acid sequences, PARP homologues and antibodies are useful for analytic detection of PARP homologues and for identifying PARP effectors or binding partners, as well as for determining their effectiveness. PARP-binding partners are useful for the diagnosis or therapy of a disease condition, which is the result of a PARP protein, especially an energy deficiency, which may comprise tissue damage from cell death following necrosis or apoptosis. The disease condition may be chosen from a neurodegenerative illness, or sepsis or ischemic tissue damage, in particular neurotoxic disturbances, etc. This sequence represents the murine PARP1 protein used in the method of the invention.

Query Match 99.6%; Score 2776.5; DB 21; Length 533;  
 Best Local Similarity 99.1%; Pred. No. 5.7e-238;  
 Matches 528; Conservative 0; Mismatches 0; Indels 5; Gaps 1;

QY 1 MAPKRASVQTEGSKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIQV 60  
 DB 1 mapkrasvqtegskqrogteedsfrstaealraapadnrvirvdpscfgrnpgiqv 60  
 QY 61 HEDYDCTLNQNTGNNGNNFYIIQLLEGSFRFCWNRWGRVGEVGSKMNHFTCLEDAKK 120  
 DB 61 hedydctlnqntgnngnnfyiqlleegsfrfcwncwgrvgrvgevggskmnhftcleddakk 120  
 QY 121 DFKKKFEKTKNWEERDRVAPQNKYTLIEVQGEAESQAVVK---VDSGPVRTVVK 175  
 DB 121 dfkkkfwetkknweerdrrvapgnytlievqgeaesgeavvkvdsqdvsgpvrvtvk 180  
 QY 176 PCSLDPATONLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQOIARGFEALAEAMK 235  
 DB 181 pcsldpatnllitnifskemfnamtlmnlvdkmpltgkltqoiargfealealeeamk 240  
 QY 236 NPTGQGSLEELSSCFYVIPHNFGRSRPPPIINSPDVLOAKKDMLLVLADIETLAQTQA 295  
 DB 241 nptgqgsleelsscfyvipnfngrsrpppnsdpvlqakkdmlvlvadielaqtlqaa 300  
 QY 296 PGEEEEKVEEVPHPDLRDYQLLRCOLQLDSESEYKATQTYLKTGNSYRCNLRHVVK 355  
 DB 301 pgeeeekveevphpldrdyqlrrcqlldsgeseykaitylktgnsyrcpnlrhvkw 360







Db 301 v-seqektveevphldrdyqlkqlqldsgapeykvigtyleqtgsnhrcptlqhiw 359  
 QY 355 KVNREGEDRFOAHSKLGNRRLLWHTGNVAVVAAITSLGRIMPHSGRGVKGIFYASEN 414  
 Db 360 kvngegedrfqahsklgnrrllwhtgtmavvaaltsglrmpshggrvkgkgyfasen 419  
 QY 415 SKSAGYVTTMHCGRHGVGMFLGELVALGKEHHITIDDPKSLSPPPGFDVSIARGOTEPDP 474  
 Db 420 sksagylvlmcgahgvymflgevalgrehhtndpslkspppgfsvlarihtepdp 479  
 QY 475 AQDTELELDGQPVVVPQPPVOCSFSSFSQSEYLYIKESQCRRLRYLLEHL 528  
 Db 480 tqdteleldgqgvvpgqpvcpsfssstfsqseylyiqesqcrlyllehl 533

RESULT 6  
 AAY68834  
 ID AAY68834 standard; Protein; 653 AA.  
 XX  
 AC AAY68834;  
 XX  
 DT 16-MAY-2000 (first entry)  
 XX  
 DE A poly(ADP-ribose) polymerase NAP protein of Zea mays.  
 XX  
 KW NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;  
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;  
 KW pest; drought; heat; fungi; nematode; seed-shatter.  
 XX  
 OS Zea mays.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..159 /note- "these residues are specifically claimed in  
 FT claim 18"  
 XX  
 PN WO200004173-A1.  
 XX  
 PD 27-JAN-2000.  
 XX  
 PF 12-JUL-1999; 99WO-EP04940.  
 XX  
 PR 17-JUL-1998; 98US-0118276.  
 XX  
 PA (PLB2 ) PLANT GENETIC SYSTEMS NV.  
 XX  
 PI Babyichuk E, Kushnir S, De Block M;  
 XX  
 DR WPI; 2000-182436/16.  
 DR N-PSDB; AAZ60616.  
 XX  
 PT Modulating cell death, growth and stress resistance in eukaryotes,  
 PT specifically plants, used, e.g. to impart fungus or nematode resistance  
 XX  
 PS Claim 18; Page 92-95; 126pp; English.  
 XX  
 CC The present sequence represents the NAP protein of Zea mays. This  
 CC protein is a poly(ADP-ribose) polymerase (PARP) protein (also known as  
 CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed  
 CC cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide  
 CC sequences can be used for modulation of programmed cell death in  
 CC eukaryotic cells. The method is used, specifically in plants, to induce,  
 CC or protect against, programmed cell death, depending on the extent to  
 CC which PARP activity is reduced. Reducing expression of endogenous NAP  
 CC class PARP only is also used to modulate programmed cell death, to  
 CC increase growth rate and to produce plant cells that are more tolerant  
 CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,  
 CC etc., or during transformation). Particular applications are generation  
 CC of plants that are resistant to fungi or nematodes; are male or female  
 CC sterile; or have better seed-shatter properties. The methods are also  
 CC used to improve growth of transformed plant cells (and derived calli or  
 CC complete plants).

XX Sequence 653 AA;  
 SQ  
 Query Match 25.8%; Score 719.5; DB 21; Length 653;  
 Best Local Similarity 35.3%; Pred. No. 6.7e-55;  
 Matches 190; Conservative 91; Mismatches 205; Indels 53; Gaps 17;  
 QY 13 GSKQRQRTGEEDSEFRS-----TAALRAAPADNRV---IRVDPSCPSFRNIGIQVHED-Y 64  
 Db 138 gakevikggdeevvkekmtatkkaavldhipdhikvnyhv-----lqvgdely 190  
 QY 65 DCTLNQINIGNNKFKYIIQLLEE--GSRFCWNRRWGRVGVGOSKMHHFTCLEL-AKKD 121  
 Db 191 datinqtvgnnnkfiiqiivlesdaggsfmvnyrvgvrgqdklbgpsptrdqaiye 250  
 QY 122 FKKFWERTKKNWEERDFVAQPNKYTLIEVO-GEAESQEAIVKVDSPGVTVVVKPCSLD 180  
 Db 251 fegkfhnnktnhwsdrknfkcyakkytwlemdygeteke-----iekgsitdqiketkle 305  
 QY 181 PATQNLTINIEFKEMFNAMTLMNLDYKKMPLGKLTQQIARGFEALAEAMKNPTGD 240  
 Db 306 triaqfislcnismmqgmveigynaekplgkirkatilkgyhvkrisdvsk--ad 363  
 QY 241 GQSLSEELSSCFYTVIPNFG--RSRPPPIPSDVLOAKKMDLLVLADIETLAQTLOAAPGE 298  
 Db 364 rrhleqtgefytviphdfgrkmrefiidtpqklkaklenvealgeleiatkl----- 418  
 QY 299 EEKVEEVPHPDLRDYQLLRQQLDLSGESEYKAIOYTLKOT---GNSYRCPNLRHVWK 355  
 Db 419 -eddsdqddplyarykqlhcdftpleadsdeysmiksylnrthgkthsgytdvdivl 477  
 QY 356 VNREGEDRFOAHSKLGNRRLLWHTGNVAVVAAITSLGRIMPH-----SGRGVKGIFYA 411  
 Db 478 vsrhgeterfqkfastrnmllwhgsrlsnwagilsglriappeavtygmfgkgyfa 537  
 QY 412 SENSKSAGYVTTMHCGRHGV---GYMFLGELVALGKEHHITIDDPKSLSPPPGFDVSIARG 468  
 Db 538 dmfsksany-----cyaseacrsgvlllcevalgdmellnadydannlpkgrskrgv 592  
 QY 469 QTEPDPAODIELELDGQPVVVPQPPVOCSFSSFSQSEYLYIKESQCRRLRYLLEIH 527  
 Db 593 qtapnm---veskvaddgvvvplgepkqeps-krggillyneyivynvdqirmryvlhvn 647

RESULT 7  
 AAR99642  
 ID AAR99642 standard; Protein; 1013 AA.  
 XX  
 AC AAR99642;  
 XX  
 DT 31-OCT-1996 (first entry)  
 XX  
 DE Poly(ADP-ribose) polymerase contg. DNA-binding domain.  
 XX  
 KW PARP; DNA-binding domain; poly(ADP-ribose) polymerase; gene therapy;  
 KW tumour treatment; DNA repair; over-expression.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9618737-A2.  
 XX  
 PD 20-JUN-1996.  
 XX  
 PF 15-DEC-1995; 95WO-DE01817.  
 XX  
 PR 16-DEC-1994; 94DE-4444949.  
 XX  
 PA (DEKR-) DEUT KREBSFORSCHUNGSZENTRUM.  
 XX  
 PI Buerkle A, Kuepper J, Zur Hausen H;  
 XX  
 DR WPI; 1996-300654/30.





PT inhibits human PARP -

XX Example 13; Page 97-101; 168pp; English.

XX The invention relates to antisense oligonucleotides targeted to human  
CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, DNA repair, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARP in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARP especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
CC arthritis) and diabetes. The present sequence is a PARP protein,  
CC the cDNA encoding which was used to design the antisense  
CC oligonucleotides.

XX Sequence 1014 AA;

Query Match 25.2%; Score 702.5; DB 22; Length 1014;  
Best Local Similarity 33.8%; Pred. No. 4.3e-53;  
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

QY 1 MAPKKASVOTEGSKKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIOV 60  
Db 493 vaprgksaga---alsksksgvkeeginksekrmkltkggaavdpdsglehsahvlegk 549  
QY 61 HEDYDCTLNQTNIGNNNNKFFYIIQLLEEG--SRFFCWNRMGRVGEV--GQSKMNHFTCTLED 117  
Db 550 gkvfsatlgldivkgtnsykqlleddkenrywifswgrvtgvisnklempsked 609  
QY 118 AKDFFKKFWEKTKNWEERDRFVAQPNKYTLIEVOGEAESQEAIVK--VDSGPVVRTVVK 175  
Db 610 aieqfmklyeektgnawhskn-ftkypkkypleid-ygqdeavkkltnvpgtksklpk 667  
QY 176 PCSLDPATQNLITNIFSKEMFNKAMTLMNLDVKMPLGKLTQKQIARGFEALEALEAMK 235  
Db 668 p-----vqdlkmifvesmkkamveyeidlqkmpigkiskrqigaaysilsevgqavs 721  
QY 236 NPTGDGOSLEELSCFVTYVPHNFRSRPPPIINSPDVLOAKMDMLVLADIELAQTLOAA 295  
Db 781 qssdsqil-dlnsrftllphdgmkkppllnnadsqvakvemdlnlldieavysllrg 835  
QY 353 VWKYNREGEDRFOAHSKLGNRLRHGTVNVAVALTSGLRIMPH---SGSRVKGKI 408  
Db 836 ifkiergecqrykpfqklnhrrllwhgstrtnfagilsgqlriapeapvtgymfgkgi 895  
QY 409 YFASENSKSAGYVTMHCGHQGVYMFGLGEVALGKEHHITIDDPDLKSPPPGDSVIARG 468  
Db 896 yfadmvsksanyhtsq--gdpigllilgevalgnmyelk-hashisrlpkghsvkglg 952  
QY 469 QTEPDPAQDLELEDGQPVVVGPPVOCPSFKSSFSQSELYLYKESQCLRLVLEI 526  
Db 953 kttdpsanl--sidg--vdvpiigtgis-sgvidtsillyneyivydiagvnlkyllki 1005

RESULT 11

AAB66296

ID AAB66296 standard; Protein; 1014 AA.

XX AAB66296;

XX AAB66296;

XX 05-APR-2001 (first entry)

XX Human tankyrase2 related protein sequence SEQ ID NO: 137.

KW Human; tankyrase2; TANK2; TRF1; telomere; cancer; neoplasm; aging;  
XX Inflammatory disorder.

OS Homo sapiens.

XX WO200100849-A1.

XX 04-JAN-2001.

XX 28-JUN-2000; 2000WO-US17827.

XX 29-JUN-1999; 99US-0141582.

XX (ICOS-) ICOS CORP.

XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;

XX WPI; 2001-102896/11.

XX N-PSDB; AAF63954.

XX New tankyrase2 polypeptides, useful for treating conditions mediated by  
PT poly(adenosine diphosphate-ribose) polymerase activity e.g. cancers,  
PT inflammatory and autoimmune disorders -  
XX Example 2; Page 207-210; 242pp; English.

XX The present invention provides the protein and coding sequence for the  
CC human tankyrase2 protein. This is found in two different versions,  
CC designated TANK2-LONG and TANK2-SHORT. Tankyrase2 has  
CC polyADP-ribosylation activity and is involved in the modification of  
CC TRF1, which is a telomere-specific binding protein. The regulation of  
CC telomere length, in which TRF1 has a role, is linked to ageing and  
CC cancer. The sequences are useful in the treatment of cancers and  
CC inflammatory disorders.

XX Sequence 1014 AA;

Query Match 25.2%; Score 702.5; DB 22; Length 1014;  
Best Local Similarity 33.8%; Pred. No. 4.3e-53;  
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

QY 1 MAPKKASVOTEGSKKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFGRNPGIOV 60  
Db 493 vaprgksaga---alsksksgvkeeginksekrmkltkggaavdpdsglehsahvlegk 549  
QY 61 HEDYDCTLNQTNIGNNNNKFFYIIQLLEEG--SRFFCWNRMGRVGEV--GQSKMNHFTCTLED 117  
Db 550 gkvfsatlgldivkgtnsykqlleddkenrywifswgrvtgvisnklempsked 609  
QY 118 AKDFFKKFWEKTKNWEERDRFVAQPNKYTLIEVOGEAESQEAIVK--VDSGPVVRTVVK 175  
Db 610 aieqfmklyeektgnawhskn-ftkypkkypleid-ygqdeavkkltnvpgtksklpk 667  
QY 176 PCSLDPATQNLITNIFSKEMFNKAMTLMNLDVKMPLGKLTQKQIARGFEALEALEAMK 235  
Db 668 p-----vqdlkmifvesmkkamveyeidlqkmpigkiskrqigaaysilsevgqavs 721  
QY 236 NPTGDGOSLEELSCFVTYVPHNFRSRPPPIINSPDVLOAKMDMLVLADIELAQTLOAA 295  
Db 722 qssdsqil-dlnsrftllphdgmkkppllnnadsqvakvemdlnlldieavysllrg 780  
QY 296 PGEEEEKVEEVPHPDLDRDYQLLRQQLDLSGSEYKAIQTYLKQT--GNSYRCPNLR--H 352  
Db 781 qssdsqil-dlnsrftllphdgmkkppllnnadsqvakvemdlnlldieavysllrg 835  
QY 353 VWKYNREGEDRFOAHSKLGNRLRHGTVNVAVALTSGLRIMPH---SGSRVKGKI 408  
Db 836 ifkiergecqrykpfqklnhrrllwhgstrtnfagilsgqlriapeapvtgymfgkgi 895  
QY 409 YFASENSKSAGYVTMHCGHQGVYMFGLGEVALGKEHHITIDDPDLKSPPPGDSVIARG 468  
Db 896 yfadmvsksanyhtsq--gdpigllilgevalgnmyelk-hashisrlpkghsvkglg 952



CC comprising an shbPARS2 nucleic acid sequence; the recombinant expression  
CC of shbPARS2; and an antibody specific for shbPARS2. shbPARS2 proteins  
CC and nucleotides are useful as vaccines for inducing an immunological  
CC response in a mammal. The shbPARS2 protein is useful for identifying  
CC compounds which inhibit or stimulate its activity or expression level.  
CC Such agonists and antagonists of shbPARS2 are useful for treating human  
CC diseases including ischaemia and ischaemic tissue injury (e.g., cerebral  
CC and cardiac ischaemia, myocardial infarction, stroke), inflammation,  
CC autoimmune disease (e.g. diabetes, multiple sclerosis) and  
CC neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's  
CC disease). shbPARS2 nucleic acids are useful as diagnostic reagents for  
CC detecting mutations in the associated gene; as hybridisation probes  
CC to isolate full-length shbPARS2 cDNAs and shbPARS2 genomic clones; and  
CC for chromosome localisation studies. The shbPARS2 protein is also useful  
CC as an immunogen to produce antibodies for therapeutic use. shbPARS2  
CC proteins, nucleotides and antibodies are also useful in screening methods  
CC for detecting the effect of added compounds on the production of mRNA and  
CC protein in cells.

XX  
XX Sequence 521 AA;

Query Match 25.1%; Score 700.5; DB 22; Length 521;  
Best Local Similarity 34.6%; Pred. No. 2.3e-53;  
Matches 185; Conservative 92; Mismatches 212; Indels 45; Gaps 17;

QY 13 GSKKQROGTBEEDSFRSTAEALRA-APADNRVIRVDPSCPFSSRNPQ---IQVHEDYDCTL 68  
Db 5 gkankdrtdedkqdesvkalllkgkap-----vdpectakvgkvhvcegnndvdyvml 57  
QY 69 NQTNIGNNNKFFYIIQLLEGRS--FFCWNRRWGRVGEVQSKMNHFTC---LEDAKKDFX 123  
Db 58 nqtnlgnnnkyyliqleddaqnrfsvmwrvgvkgmghsl--vacsgnlmkakeifq 115  
QY 124 KFEWETKKNWEERDRFVAQPNKYTLIEVQGEAESEAVVVDGSPVTVVVKPCS-LDPA 182  
Db 116 kfkldtknnwedrekfepvgkydmlqmdyatntqdeetkkeslksplkpesqlidr 175  
QY 183 TONLITNIFSKEMFNKAMTLMNLDVKKMPLGLTKQOIARGFPALEALEAMKNPTGD-G 241  
Db 176 vqelikicnvqameememkynntkkgaplgltvqagikagyslkkiedcir--agqhg 233  
QY 242 QSLLELSSCFYTVIPNFGSRPPPIINSPDVLOAKKMDLLVLADIETLAQTLQAAPGEEEE 301  
Db 234 ralmeacnefytriphdfgrtptlirtqksekiglealgieiaikl-----vkt 287  
QY 302 KVEEVPHPDLDRDYQLLRCOLQLDLSGESEYKAIQTVLKQT---GNSYRCPNLRHVKNR 358  
Db 288 eiqspehpldqhrynlhcalrpdhesyefkvisqylqsthapthdsytdmtlidlfevek 347  
QY 359 EGEGDRFQAHSKILGNRRLLWHGNTNVAVVAAILTSGLRIMPH----SGGRVKGKIYFASEN 414  
Db 348 dgekeaf--edlhnrmllwhgrmsnwqilshgriapeapitgymfgkgyfadms 405  
QY 415 SKSAGYVTHMCHGQGVQYMFLEVALGKEHHHTIDDPKSLKPPPGFSDVIARGQTEPDP 474  
Db 406 skeanyefasrl--kntglllsevalgqcnleapkaegllqgkstkiglkmapss 463  
QY 475 AQDIELELDGQPVVPGPPVQCPKSSSF--SQSEYLLYKESQCELRLLLEI 526  
Db 464 ahiv--clngs--tvplgpasdtgilnpdgytlnyneyivnpgnvmryllkv 513

## RESULT 14

AY51174  
ID AAY51174 standard; Protein; 570 AA.

XX

AC AAY51174;

XX

DT 31-MAR-2000 (first entry)

XX Human brain PARP2 protein.

DE

XX

KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage.

XX Homo sapiens.

XX W09964572-A2.

XX 16-DEC-1999.

XX 04-JUN-1999; 99WO-EP03889.

XX 05-JUN-1998; 98DE-1025213.

XX 01-MAR-1999; 99DE-1008837.

XX (BADI ) BASF AG.

XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;

XX WPI; 2000-087218/07.

XX N-PSDB; AA244287.

XX Novel genes and proteins, antibodies and binding partners useful in  
XX diagnosis and therapy of energy deficiency associated disease  
XX conditions -

XX Claim 4; Page 52-54; 96pp; German.

XX This invention describes novel human and murine poly(ADP-ribose)  
XX polymerase (PARP) homologues, which are characterised by an amino acid  
XX sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
XX sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MHX<sub>2</sub>C(I). The nucleic acid  
XX sequences, PARP homologues and antibodies are useful for analytic  
XX detection of PARP homologues and for identifying PARP effectors or  
XX binding partners, as well as for determining their effectiveness.  
XX PARP-binding partners are useful for the diagnosis or therapy of a  
XX disease condition, which is the result of a PARP protein, especially an  
XX energy deficiency, which may comprise tissue damage from cell death  
XX following necrosis or apoptosis. The disease condition may be chosen  
XX from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
XX in particular neurotoxic disturbances, etc. This sequence represents the  
XX human PARP2 protein used in the method of the invention.

XX Sequence 570 AA;

Query Match 25.1%; Score 700.5; DB 21; Length 570;  
Best Local Similarity 34.6%; Pred. No. 2.7e-53;  
Matches 185; Conservative 92; Mismatches 212; Indels 45; Gaps 17;

QY 13 GSKKQROGTBEEDSFRSTAEALRA-APADNRVIRVDPSCPFSSRNPQ---IQVHEDYDCTL 68  
Db 54 gkankdrtdedkqdesvkalllkgkap-----vdpectakvgkvhvcegnndvdyvml 106  
QY 69 NQTNIGNNNKFFYIIQLLEGRS--FFCWNRRWGRVGEVQSKMNHFTC---LEDAKKDFX 123  
Db 107 nqtnlgnnnkyyliqleddaqnrfsvmwrvgvkgmghsl--vacsgnlmkakeifq 164  
QY 124 KFEWETKKNWEERDRFVAQPNKYTLIEVQGEAESEAVVVDGSPVTVVVKPCS-LDPA 182  
Db 165 kfkldtknnwedrekfepvgkydmlqmdyatntqdeetkkeslksplkpesqlidr 224  
QY 183 TONLITNIFSKEMFNKAMTLMNLDVKKMPLGLTKQOIARGFPALEALEAMKNPTGD-G 241  
Db 225 vqelikicnvqameememkynntkkgaplgltvqagikagyslkkiedcir--agqhg 282  
QY 242 QSLLELSSCFYTVIPNFGSRPPPIINSPDVLOAKKMDLLVLADIETLAQTLQAAPGEEEE 301  
Db 283 ralmeacnefytriphdfgrtptlirtqksekiglealgieiaikl-----vkt 336  
QY 302 KVEEVPHPDLDRDYQLLRCOLQLDLSGESEYKAIQTVLKQT---GNSYRCPNLRHVKNR 358  
Db 337 eiqspehpldqhrynlhcalrpdhesyefkvisqylqsthapthdsytdmtlidlfevek 396



**THIS PAGE BLANK (USPTO)**

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:58:21 ; Search time 46.79 Seconds  
(without alignments)  
275.630 Million cell updates/sec

Title: US-09-701-586B-10  
Perfect score: 2789  
Sequence: 1 MAPKKASVQTGSKKQROG.....EVLTKESQRLRYLLEIHL 528

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	707.5	25.4	1013	4	US-08-860-886-2
2	702.5	25.2	1014	4	US-08-078-347A-3
3	143	5.1	1327	4	US-09-196-387-2
4	106.5	3.8	635	4	US-09-046-992-2
5	104	3.7	2285	4	US-09-308-375-2
6	103.5	3.7	613	1	US-08-405-615-1
7	103.5	3.7	613	2	US-08-461-234-1
8	103.5	3.7	613	2	US-08-463-480-1
9	103.5	3.7	614	1	US-08-225-224-1
10	103.5	3.7	614	3	US-08-722-258-1
11	103.5	3.7	614	5	PCT-US95-04468-1
12	103.5	3.7	638	3	US-09-047-148-2
13	103	3.7	462	2	US-08-484-438-42
14	102.5	3.7	3075	2	US-08-460-309-5
15	102.5	3.7	3075	2	US-08-125-077-5
16	101	3.6	622	2	US-08-356-786-16
17	100.5	3.6	420	1	US-08-391-259-7
18	100.5	3.6	420	2	US-08-839-425-7
19	100.5	3.6	2482	1	US-08-328-254-6
20	99	3.5	579	1	US-08-126-564A-31
21	99	3.5	579	5	PCT-US94-09143-31
22	98	3.5	889	4	US-09-336-447A-15
23	97	3.5	662	1	US-08-224-657-88
24	97	3.5	662	4	US-09-354-138-88
25	96.5	3.5	497	2	US-08-511-485-4
26	96.5	3.5	497	3	US-09-212-971-4
27	96.5	3.5	497	4	US-08-800-929A-4

28	96.5	3.5	497	4	US-09-617-053A-4	Sequence 4, Appli
29	96	3.4	163	1	US-08-044-618-6	Sequence 6, Appli
30	96	3.4	637	1	US-08-235-838-14	Sequence 14, Appl
31	96	3.4	637	2	US-08-465-473B-14	Sequence 14, Appl
32	95.5	3.4	396	4	US-09-046-992-4	Sequence 4, Appli
33	95.5	3.4	420	1	US-08-391-259-2	Sequence 2, Appli
34	95.5	3.4	420	1	US-08-391-259-2	Sequence 2, Appli
35	95.5	3.4	420	2	US-08-839-425-2	Sequence 11, Appl
36	95.5	3.4	420	2	US-08-839-425-11	Sequence 2, Appli
37	95.5	3.4	599	1	US-08-463-163-3	Sequence 11, Appl
38	95.5	3.4	637	1	US-08-235-838-16	Sequence 3, Appli
39	95.5	3.4	637	2	US-08-465-473B-16	Sequence 16, Appl
40	95.5	3.4	1128	4	US-08-923-992A-6	Sequence 6, Appli
41	95.5	3.4	1164	4	US-08-923-992A-2	Sequence 2, Appli
42	95.5	3.4	3248	1	US-08-353-700-1	Sequence 1, Appli
43	95.5	3.4	3248	5	PCT-US95-16216-1	Sequence 1, Appli
44	95	3.4	920	1	US-08-451-715A-2	Sequence 2, Appli
45	94.5	3.4	1098	4	US-08-923-992A-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1  
US-08-860-886-2  
; Sequence 2, Application US/08860886  
; Patent No. 6335009  
; GENERAL INFORMATION:  
; APPLICANT: Burkli, Alexander  
; APPLICANT: Zur Hausen, Harald  
; APPLICANT: Jan-Heiner, Kupper  
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE  
; TITLE OF INVENTION: IN GENE THERAPY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds, LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/860,886  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 8484-0028-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-08-860-886-2

Query Match 25.4%; Score 707.5; DB 4; Length 1013;

Best Local Similarity 34.0%; Pred. No. 3.3e-60;  
Matches 183; Conservative 101; Mismatches 216; Indels 39; Gaps 17;

QY 1 MAPKRASVQTGSKKORQCTEEDSFRSTAEALRAAPADNRVIRVDPSPCFSPNPGIOV 60  
Db 492 VAPRGSGA---ALSRSKQGVKEEGINKSEKRMKLTGKGAADVDPDGLGSAHVLEK 548  
QY 61 HEDYDCTLNQTNIGNNNNKFIYIQLLEEG--SRFFCNRWRGVGEV--GOSKMNHFTCLE 117  
Db 549 GKVSATLGLVDIVKGTNSYKQLLEDKKNRYWIFRSWGRVGTIGSKNKLQMPSPKED 608  
QY 118 AKDFFKKFWEKTKNWEEDRFVAQPNKYTLIEVOGEAESQAVVK--VDSGPVPTVVK 175  
Db 609 AIEHFMKLYEKTGNMWSKN-FTKYPKFPLEID-YGQDEEAVKLTVPNPGTKSKLPK 666  
QY 176 PCSLDPATQNLINIFSKEMFNKNTLMNLDVKKMPLGKLTQKQIARGFEALEEAMK 235  
Db 667 P-----VDQLIKMIFDVESSMKAMVEYEDLQKMPGLSKRQIQAYSILSEVOQAVS 720  
QY 236 NPTGDSLEELSCFVTIVPHNFRSRPPINSPOVLAQKMDMLVLADIELAQTLOAA 295  
Db 721 QGSDDSOIL-DLNRFTYLLPHDFGKKPPLNNADSVQAKVEMLDNLLDIEVAYSLLRG 779  
QY 296 PGEEKEVEEVPPLDRDYQLLRCQLQDLSGESEYKAIQTYLKQ----TGNSYRCPNLR 351  
Db 780 GSDSSK-----DPIDVNEKLTIDKVVDRDSEAEIIRKYVKNTHATTNAYDL-EVI 833  
QY 352 HVKVNREGEDRQAHSKLGNRRLLWHGTVNAVVAAILTSLGRIMPH-----SGGRVGK 407  
Db 834 DIPKIERGECQYKPKQLHNRLLWHGSRRTTNFAGILSQGLRIAPPEAPVGYMFGKI 893  
QY 408 IYFASENKSAGYVVTMHCCHGVYMFGLGEVALGKHEHITIDDPKSLSPPPGDFSVIAR 467  
Db 894 IYFADVMSKSNYCHTSQ--GDPGILLGELVALGNMYELK-HASHISKLPGKHVSKGL 950  
QY 468 QTEPDPAQDIELELDQPVVPPVQCPSPKSSFSQSEYLYIYKESQCRLYLLEI 526  
Db 951 KTTDPSPANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEIYVDTAQNVLKYLKL 1004

RESULT 2  
US-09-701-586b-3  
; Sequence 3, Application US/09078347A  
; Patent No. 6132968  
; GENERAL INFORMATION:  
; APPLICANT: Le, Xiao-Chun  
; APPLICANT: Weinfield, Michael  
; APPLICANT: Xing, James Z.  
; TITLE OF INVENTION: Methods for Quantitating Low Level  
; TITLE OF INVENTION: Modifications of Nucleotide Sequences  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,347A  
; FILING DATE: 13-MAY-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.  
; REGISTRATION NUMBER: 32,837  
; REFERENCE/DOCKET NUMBER: UALB-03283  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410

TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1014 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: not relevant  
; TOPOLOGY: unknown  
; MOLECULE TYPE: protein  
US-09-701-586b-3

Query Match 25.2%; Score 702.5; DB 4; Length 1014;  
Best Local Similarity 33.8%; Pred. No. 1e-59;  
Matches 182; Conservative 102; Mismatches 217; Indels 37; Gaps 17;

QY 1 MAPKRASVQTGSKKORQCTEEDSFRSTAEALRAAPADNRVIRVDPSPCFSPNPGIOV 60  
Db 493 VAPRGSGA---ALSRSKQGVKEEGINKSEKRMKLTGKGAADVDPDGLGSAHVLEK 549  
QY 61 HEDYDCTLNQTNIGNNNNKFIYIQLLEEG--SRFFCNRWRGVGEV--GOSKMNHFTCLE 117  
Db 550 GKVSATLGLVDIVKGTNSYKQLLEDKKNRYWIFRSWGRVGTIGSKNKLQMPSPKED 609  
QY 118 AKDFFKKFWEKTKNWEEDRFVAQPNKYTLIEVOGEAESQAVVK--VDSGPVPTVVK 175  
Db 610 AIEHFMKLYEKTGNMWSKN-FTKYPKFPLEID-YGQDEEAVKLTVPNPGTKSKLPK 667  
QY 176 PCSLDPATQNLINIFSKEMFNKNTLMNLDVKKMPLGKLTQKQIARGFEALEEAMK 235  
Db 668 P-----VDQLIKMIFDVESSMKAMVEYEDLQKMPGLSKRQIQAYSILSEVOQAVS 721  
QY 236 NPTGDSLEELSCFVTIVPHNFRSRPPINSPOVLAQKMDMLVLADIELAQTLOAA 295  
Db 722 QGSDDSOIL-DLNRFTYLLPHDFGKKPPLNNADSVQAKVEMLDNLLDIEVAYSLLRG 780  
QY 296 PGEEKEVEEVPPLDRDYQLLRCQLQDLSGESEYKAIQTYLKQ--GNSYRCPNLR--H 352  
Db 781 GSDSSK-----DPIDVNEKLTIDKVVDRDSEAEIIRKYVKNTHATTNAYDL-EVI 835  
QY 353 HVKVNREGEDRQAHSKLGNRRLLWHGTVNAVVAAILTSLGRIMPH-----SGGRVGK 408  
Db 836 IFKIERGECQYKPKQLHNRLLWHGSRRTTNFAGILSQGLRIAPPEAPVGYMFGKI 895  
QY 409 IYFASENKSAGYVVTMHCCHGVYMFGLGEVALGKHEHITIDDPKSLSPPPGDFSVIAR 468  
Db 896 IYFADVMSKSNYCHTSQ--GDPGILLGELVALGNMYELK-HASHISKLPGKHVSKGL 952  
QY 469 QTEPDPAQDIELELDQPVVPPVQCPSPKSSFSQSEYLYIYKESQCRLYLLEI 526  
Db 953 KTTDPSPANI--SLDG--VDVPLGTGIS-SGVNDTSLLYNEIYVDTAQNVLKYLKL 1005

RESULT 3  
US-09-196-387-2  
; Sequence 2, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor  
; CITY: Hackensack  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

```
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196.387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-196-387-2
```

```
Query Match 5.18; Score 143; DB 4; Length 1327;
Best Local Similarity 23.98; Pred. No. 9.5e-05;
Matches 71; Conservative 41; Mismatches 107; Indels 78; Gaps 13;

Qy 271 DVLAQKMLLVLDI---ELAQTLQAAPGEEKEKVEVPHPPL-----DRDYQLLRCLQL 322
   | : : : | | | | | : : : | : : : | : : : | : : : | : : : |
Db 1045 DIFETEITLDVLDAMGHEELKEIGINAYGHRKLIKVERLLGGQGTNPYLTFHCYNQ 1104

Qy 323 ---LLDSG--ESEYKAIQTYLKQT-----GNSVRCPNLRHVWVKNREGEGRF-- 365
   | : : : | | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1105 GTILDLAPEDKEYQSVSEEMQSTIREHRDGNAGGIENRYNVIQKVVNKKLRERFCH 1164

Qy 366 -----QAHSKLGNNRLAHGNTNVAVVAALITSL-RLMPHSGGRVKGKIYFASENSKSA 418
   | : : : | | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1165 RQKEVSENNHNNHRLMFLHGS--PFINAIHKGFDERHAYIGMGFAGIYFASNSSKSN 1222

Qy 419 GYVTM-----RCGGHQ-----VGYMFLGEVAGLCGEHHTIDDPSLKSPPPGDFSVIAR 467
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 1223 QYVYIGGGTGCPTHKDRSCYICHRQLFCRVTLGKS-FLQFSIMKMAHAPPGHHSVI-- 1279

Qy 468 GQTEPDPAQDIELDGQPVVVGQPPVQCPSPFKSSFSQSEYLIYKESQCLRLRYLL 524
   | : : | : : | : : | : : | : : | : : | : : | : : | : : |
Db 1280 -----GRPSV-----NGLAYAEYVIYRGEQAYPEYLI 1306
```

```
RESULT 4
US-09-046-992-2
; Sequence 2, Application US/09046992
; Patent No. 6140066
; GENERAL INFORMATION:
; APPLICANT: Lorberbaum-Galski, Haya
; APPLICANT: Yarkoni, Shai
; APPLICANT: Ben-Yehudah, Ahmi
; TITLE OF INVENTION: METHODS OF CANCER DIAGNOSIS
; TITLE OF INVENTION: USING A CHIMERIC TOXIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds, LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: NY
; COUNTRY: USA
; ZIP: 10036-2811
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
```

```
; OPERATING SYSTEM: Windows
; SOFTWARE: FastSeq for Windows Version 2.0b
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/046.992
; FILING DATE: 24-MAR-1998
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Poissant, Brian M
; REGISTRATION NUMBER: 28,462
; REFERENCE/DOCKET NUMBER: 9457-0013-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-493-4935
; TELEFAX: 650-493-5556
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 635 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; US-09-046-992-2

Query Match 3.88; Score 106.5; DB 4; Length 635;
Best Local Similarity 19.08; Pred. No. 0.11;
Matches 79; Conservative 55; Mismatches 154; Indels 127; Gaps 17;

Qy 148 TLIEVQGAESQEAQVVKVDSGPVRTVVKP-----CSDLPATQNLITNFFSKEM 195
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 195 SVVMAQNPQPREKRWSEWASGVKVLCLDPLDGVVNYLAAQRCNLDLTWEGKIYRVLAGNP 254

Qy 196 FKNAMTLMNLDVK-----KMPGLKLTQKQIARGFEALE-- 228
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 255 AKH-----DLDIKPTVISEELEFPPEGGSIAALTAHQACHLPLETTRHRQPRGWQLEQC 309

Qy 229 -----ALEEAMKNPTGQSGSLEELSSCFYVIPHNFORSRPP 265
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 310 GYPVQLRVLYLAARLSWNQDVIRNALASP-GSGGDLGEAI-----REQPE 356

Qy 266 PINSDPVLQAQKMLLVLDIETLAQTLQAAPGEEKEKVEVPHPPLDRDYQLLRCLQL-- 323
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 357 -----QAR--LALTAAAESERFVQGTGNDAGA-----ANADVSLTCTPVAAGE 400

Qy 324 ---LDSG-----ESEYKAIQTYLKQTSYRCPNLRHVWVKNREGEGRFQAHSKLGNNR 375
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 401 CAGPADSGDALLEANYPTGAEFGLDGDGVSTFTRGTQNTVTER-----LLQAHRQLEERG 455

Qy 376 LL---WHGTNVAVVAALITSLRIMPHSGGRVKGKIYFASENSKSAGYVTTMHCGGHQVG 432
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 456 YVFGYHGTFLEAAQSIYFVGGRVARSQDLDAIWRGFYIAGDPALAYGAQDQEPDAR--- 512

Qy 433 YMFELGEVALGKEHHTIDDPSLKSPPPGDFSDVIARGQTEPOPAQDIELELDGQPV 487
   | : : : | : : : | : : : | : : : | : : : | : : : | : : : |
Db 513 ---GRIRNGALLRVYVPRSSL-----PGFYRT-SLTLAAPAAAGEVE-RLIGHPL 557

RESULT 5
US-09-308-375-2
; Sequence 2, Application US/09308375
; Patent No. 6300117
; GENERAL INFORMATION:
; APPLICANT: Genencor
; TITLE OF INVENTION: Proteases From Gram-Positive Organisms
; FILE REFERENCE: GC394-PCT
; CURRENT APPLICATION NUMBER: US/09/308.375
; CURRENT FILING DATE: 1999-05-14
; EARLIER APPLICATION NUMBER: EP9719636.4
; EARLIER FILING DATE: 1997-09-15
```





COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/225,224  
FILING DATE: 8-APR-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen L.  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 15280-193  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..614  
OTHER INFORMATION: /label= native-pe  
US-08-225-224-1

Query Match 3.7%; Score 103.5; DB 1; Length 614;  
Best Local Similarity 19.7%; Pred. No. 0.21;  
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;  
QY 167 SGPVRTVVKP-----CSLDPATONLITNIFSKEMFNAMTLMNLDVK----- 208  
DB 193 SGKVLCLDPLDGVYNYLAQQRCLNDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247  
QY 209 -----KMPGLKTKQOIARGFEALE----- 228  
DB 248 RLHFPEGGSLAALTAHQACHLPLETFRHRQPRGWEOLEQCGYPVQLVALYLAARLSWN 307  
QY 229 -----ALEEAMKNPTGQSQSLEELSSCFYVPHNFGSRPPPIPSDPVLOAKKMDLLVLA 284  
DB 308 QVDQVIRNALASP-GSGGDLGEAI-----REQPE-----QAR--LALTILA 344  
QY 285 DIELAQTLQAAPGEEKEVEVPHLDROYQLLRCOLQL-----LDSG-----ESEYKAI 334  
DB 345 AESERFVROGTGNDEAGA-----ANADVSLTCPVAAGECAGPADSGDALLERNYPTG 398  
QY 335 QTYLKOTGNSYRCPNLRHVYKVNREGEDRFOAHSKLGNNRLL---WHGTNVAVVAAILT 391  
DB 399 AEFLGDGDSVSFSTRGTQNTVER-----LLQAHROLEERGYYVFGYHGTFLEAAQSIVF 453  
QY 392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVYTHMCGGHQGVYMFLEVALGKEHHITIDD 451  
DB 454 GGVRARSDQLDAIWRGFYIAGDPALAYGAQDOEPDAR-----GRIRNGALLRVVPR 506  
QY 452 PSLKSPPPGDSVIARGOTEPDPAQDIELELDGPV 487  
DB 507 SSL-----PGFYRT-SUTLAAPEAAEVE-RLIHPL 536  
RESULT 10  
US-08-722-258-1  
; Sequence 1, Application US/08722258  
; Patent No. 6011002  
; GENERAL INFORMATION:  
; APPLICANT: Pastan, Ira  
; APPLICANT: Kreitman, Robert J.  
; APPLICANT: Puri, Raj K.  
; TITLE OF INVENTION: Circularly Permuted Ligands and  
; TITLE OF INVENTION: Circularly Permuted Chimeric Molecules

NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/722,258  
FILING DATE: 08-JAN-1997  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: WO PCT/US95/04468  
FILING DATE: 06-APR-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/225,224  
FILING DATE: 08-APR-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 015280-193100US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..614  
OTHER INFORMATION: /note= "native Pseudomonas exotoxin"  
OTHER INFORMATION:  
US-08-722-258-1

Query Match 3.7%; Score 103.5; DB 3; Length 614;  
Best Local Similarity 19.7%; Pred. No. 0.21;  
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;  
QY 167 SGPVRTVVKP-----CSLDPATONLITNIFSKEMFNAMTLMNLDVK----- 208  
DB 193 SGKVLCLDPLDGVYNYLAQQRCLNDDTWEGKIYRVLAGNPAKH-----DLDIKPTVISH 247  
QY 209 -----KMPGLKTKQOIARGFEALE----- 228  
DB 248 RLHFPEGGSLAALTAHQACHLPLETFRHRQPRGWEOLEQCGYPVQLVALYLAARLSWN 307  
QY 229 -----ALEEAMKNPTGQSQSLEELSSCFYVPHNFGSRPPPIPSDPVLOAKKMDLLVLA 284  
DB 308 QVDQVIRNALASP-GSGGDLGEAI-----REQPE-----QAR--LALTILA 344  
QY 285 DIELAQTLQAAPGEEKEVEVPHLDROYQLLRCOLQL-----LDSG-----ESEYKAI 334  
DB 345 AESERFVROGTGNDEAGA-----ANADVSLTCPVAAGECAGPADSGDALLERNYPTG 398  
QY 335 QTYLKOTGNSYRCPNLRHVYKVNREGEDRFOAHSKLGNNRLL---WHGTNVAVVAAILT 391  
DB 399 AEFLGDGDSVSFSTRGTQNTVER-----LLQAHROLEERGYYVFGYHGTFLEAAQSIVF 453  
QY 392 SGLRIMPHSGGRVGKGIYFASENSKSAGYVYTHMCGGHQGVYMFLEVALGKEHHITIDD 451  
DB 454 GGVRARSDQLDAIWRGFYIAGDPALAYGAQDOEPDAR-----GRIRNGALLRVVPR 506

DB 454 GGVARSQDLDAIWRFYIAGDPALAYGYAQDQEPDAR-----GRIRNGALLRVYVPR 506

Db 423 AEFLGDCGVSFSTRGTQNTWVER-----LLQAHRLQLEERGVYFVGHGTFLEAAQSVF 477  
QY 392 SGLRIMPHSGRGYKGIYFASENSKSAGYVTTMHCGGHQVGYMFLGVALGKEHHITIDD 451  
Db 478 GGYRARSQDLDAIWGRFYIAGDPAALAYGQAQDQEPDAR-----GRIRNGALLRVVPR 530  
QY 452 PSLSKPPPGDSVIARTQTEPDPAQDIELELDGQPV 487  
Db 531 SSL-----PGFYRT-SLTAAPEAAAGEVE-RLIGHPL 560

RESULT 13  
US-08-484-438-42  
; Sequence 42, Application US/08484438  
; Patent No. 5811098  
; Patent No. 5811098 5780031  
; GENERAL INFORMATION:  
; APPLICANT: Plowman, Gregory D.  
; APPLICANT: Culouscou, Jean-Michel  
; APPLICANT: Shoyab, Mohammed  
; APPLICANT: Siegall, Clay B.  
; APPLICANT: Hellstr m, Ingegerd  
; APPLICANT: Hellstr m, Karl E.  
; TITLE OF INVENTION: HER4 HUMAN RECEPTOR TYROSINE KINASE  
; NUMBER OF SEQUENCES: 42  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/484,438  
; FILING DATE: 07-JUN-1995  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/323,442  
; FILING DATE: 14-OCT-1994  
; APPLICATION NUMBER: US 08/150,704  
; FILING DATE: 10-NOV-1993  
; CLASSIFICATION: 530  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/981,165  
; FILING DATE: 24-NOV-1992  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Misrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 5624-230  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 42:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-484-438-42

Query Match 3.7%; Score 103; DB 2; Length 462;  
Best Local Similarity 21.1%; Pred. No. 0.15;  
Matches 95; Conservative 46; Mismatches 181; Indels 128; Gaps 21;  
QY 123 KKKFWETKKNWEERDFVAQPNKYTLIEVQGEAESQEAQVAVKVDGSPVRTVVRKPCSLD-- 180

Db 22 KKGKNGKNNR--RNRSHLTKCAEKETFCVNG---GECFTVKDLSNPSRYLCK--CPNEFT 76  
QY 181 -PATONLITNIFSK--EMFK-----NMTLMNLDVKKMPLGKLTQQIARGFEALE-- 228  
Db 77 GDRCONVVASFYKABELYKLMAEEGSLAALTAHQACHLPLETTFFHRRQPRGWEQLEQC 136  
QY 229 -----ALBEMKNPTGDGOSLEELSSCFYTVVPHNFGSRPP 265  
Db 137 GYPVQRVALYLAARLSWNQDVIRNALASP--GSGDLGEAI-----RQPE 183  
QY 266 PINSPDVLQAKDMLLVLADELIAQTLQAPGEEEEKVEEVEPHPLDREDYQLLRCOL-- 323  
Db 184 -----QAR--LALTAAEESERFVROGTGNDAGA-----ANADVVSLLTCPVAAGE 227  
QY 324 ----LDSG-----ESEYKAIQTYLTKQTSYRCPNLRHVWKNREGEGRDQFAHSLGNRR 375  
Db 228 CAGPADSGDALLERNYPTGAFLDGDGDSFSTRGTQNTWVER-----LLQAHRLQLEERG 282  
QY 376 LL---WHGTNVAVVAAILTSGLRIMPHSGRGVKGIGYFASENSKSAGYVTTMHCGGHQVG 432  
Db 283 YFVGYHGTFLAAQSIIVGGVRRARSODLDAIWRGFYIAGDPALAYGQAQDQEPDAR-- 339  
QY 433 YMFGLGEVALGKEHHITIDPSLSKSPGPFDSVI-----ARGQTEPDPAQDIELELD-- 483  
Db 340 ----GRIRNGALLRVYPRSSL-----PGFYRTSLTLAGEAAGEVERLIGHPLRLDAI 391  
QY 484 -----GDP-----VWVPGPP 494  
Db 392 TGPEERGGRLTETLGNWPLAERTVVIIPSAIP 421

RESULT 14  
US-08-460-309-5  
; Sequence 5, Application US/08460309  
; Patent No. 5837496  
; GENERAL INFORMATION:  
; APPLICANT: Engvall, Eva  
; APPLICANT: Leivo, Ilmo  
; TITLE OF INVENTION: Nucleic Acids Encoding Merosin, Merosin  
; TITLE OF INVENTION: Fragments and Uses Thereof  
; NUMBER OF SEQUENCES: 23  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/460,309  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/125,077  
; FILING DATE: 22-SEP-1993  
; APPLICATION NUMBER: US PCT/US 94/10730  
; FILING DATE: 21-SEP-1994  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/472,319  
; FILING DATE: 30-JAN-1990  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/919,951  
; FILING DATE: 27-JUL-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-LA 9721





Job time: 302 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:19 ; Search time 36.19 Seconds  
(without alignments)  
570.255 Million cell updates/sec

Title: US-09-701-586B-8  
Perfect score: 2813  
Sequence: 1 MAPKRKASVQTEGSKKQKG.....EYLIYKESQCLRYLLEIHL 533

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2229	79.2	533	PP03_HUMAN	Q9Y6f1 homo sapien
2	713.5	25.4	1012	PP0L_CRIGR	Q9r152 cricetus
3	708	25.2	559	PP02_MOUSE	O88554 mus musculus
4	705	25.1	1013	PP02_HUMAN	P09874 homo sapien
5	702	25.0	1015	PP0L_BOVIN	P18493 bos taurus
6	696.5	24.8	583	PP02_HUMAN	Q9ug85 homo sapien
7	694.5	24.7	1013	PP0L_RAT	P27008 rattus norv
8	691.5	24.6	1011	PP0L_CHICK	P26446 gallus gall
9	690	24.5	637	PP0L_ARATH	Q11207 arabidopsis
10	686.5	24.4	998	PP0L_XENLA	P31669 xenopus lae
11	683.5	24.3	1012	PP0L_MOUSE	P11103 mus musculus
12	645.5	22.9	996	PP0L_SARPE	Q11208 sarcophaga
13	619	22.0	994	PP0L_DROME	P35875 drosophila
14	396.5	14.1	538	Y0N4_CAEEL	Q09525 caenorhabdi
15	275	9.8	1724	PP0V_HUMAN	Q9uk83 homo sapien
16	193	6.9	135	PP0L_OGMA	Q08824 oncorhynch
17	117	4.2	2035	EVPL_MOUSE	Q9d952 mus musculus
18	106.5	3.8	638	TOXA_PSEAE	P11439 pseudomonas
19	106.5	3.8	1031	KINH_STRPU	P35978 strongyloce
20	104.5	3.7	963	KINH_HUMAN	P33176 homo sapien
21	104	3.7	1395	SP41_YEAST	P38904 saccharomyc
22	104	3.7	3210	CENF_HUMAN	P49454 homo sapien
23	102	3.6	752	CAT1_MYCFO	O08404 mycobacteri
24	101.5	3.6	1901	YCF1_TOBAC	P12222 nicotiana t
25	101	3.6	1085	RBP2_BOVIN	P48820 bos taurus
26	101	3.6	3075	LMAL_HUMAN	P25391 homo sapien
27	100.5	3.6	1574	RPOC_AQUAE	O67763 aquifex aeo
28	100	3.6	583	LAMI_XENLA	P09010 xenopus lae
29	99	3.5	495	NUSA_ECOLI	P03003 escherichia
30	99	3.5	497	BIRA_HUMAN	P98170 homo sapien
31	99	3.5	638	RPAL_EUPOC	P28363 euplotes oc
32	99	3.5	646	YE14_SCHPO	O13869 schizosacch
33	99	3.5	890	RB6K_HUMAN	O95235 homo sapien

34	98.5	3.5	583	1	T2F1_FLAOK	P14870 flavobacter
35	98.5	3.5	687	1	NNP1_DROME	Q9vjz7 drosophila
36	98	3.5	612	1	EXO2_BPT5	P11109 bacterioph
37	98	3.5	830	1	YBMA_SCHPO	O10332 schizosacch
38	98	3.5	963	1	KINH_MOUSE	Q61788 mus musculu
39	97.5	3.5	379	1	Y011_MOUSE	P11260 mus musculu
40	97.5	3.5	603	1	US26_HCMVA	P09699 human cytom
41	97	3.4	725	1	ADDB_MOUSE	Q9qy88 mus musculu
42	96.5	3.4	1005	1	E4L2_HUMAN	O43491 homo sapien
43	96.5	3.4	2453	1	NCRL_MOUSE	O60974 mus musculu
44	96.5	3.4	5327	1	ACF7_MOUSE	Q9qxx20 mus musculu
45	96	3.4	629	1	PAB2_ARATH	P42731 arabidopsis

ALIGNMENTS

RESULT	1
PP03_HUMAN	STANDARD; PRT; 533 AA.
AC	Q9Y6f1: Q9UG81: 16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DE	Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+) ADP-riboseyltransferase-3) (Poly[ADP-ribose] synthetase-3) (pADPRT-3)
DE	(hPARP-3).
GN	ADPRTL3 OR PARP3 OR ADPRT3.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=Fetal brain;
RX	MEDLINE=99263509; PubMed=10329013;
RA	Johansson M.;
RT	"A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA cloning of two novel poly(ADP-ribose) polymerase homologues.";
RL	Genomics 57:442-443(1999).
RN	[2]
RC	SEQUENCE OF 75-533 FROM N.A.
RA	TISSUE=Kidney;
RC	Ansoerge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;
RL	Submitted (May-1999) to the EMBL/GenBank/DBJ databases.
CC	-I- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor -> nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC	-I- SUBCELLULAR LOCATION: Nuclear (Potential).
CC	-I- TISSUE SPECIFICITY: Widely expressed; the highest levels are in the kidney, skeletal muscle, liver, heart and spleen; also detected in pancreas, lung, placenta, brain, leukocytes, colon, small intestine, ovary, testis, prostate and thymus.
CC	-I- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
DR	EMBL; AF083068; AAD29855.1; -
DR	EMBL; AL050034; CAB43246.1; -
DR	HSSP; P26446; 1A26.
DR	InterPro; IPR001290; PARP.
DR	InterPro; IPR004102; PARP_reg.
DR	Pfam; PF00644; PARP; 1.
DR	Pfam; PF02877; PARP; reg; 1.
DR	Transferase; Glycosyltransferase; NAD; Nuclear protein;
KW	ADP-ribosylation.
FT	DOMAIN 14 20 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
FT	CONFLICT 80 80 K -> N (IN REF. 2).



Db 493 APRKSAAPSKSK-----GLYKEGVNKKSEKMKLTKGAAVDPDGLSAHVLEKGG 548  
 QY 62 EDYDCTLNQTNIGNNNNKYYIIQLLEEG--SRFFCHNRGRVGEV-QSKMNHFTCLEDA 118  
 Db 549 KVSATLGLVIVKGVNSYKQLLEDDEKESYWIYFRSGRWGVTVIGSNKLEQMPKSKEDA 608  
 QY 119 KDKFKKFKWETKKNKEEDRFRVAOPKNYKTLTEVOGEAESQAVVKALSPQVDSGVRV 178  
 Db 609 VEHEMKLYEKTGNWHSKN-FTKYPKFPLEID-YGQDEAVKK-----LT 654  
 QY 179 VXP---CSLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQQTARGFEALEAL 235  
 Db 655 VXPGTKSLPKPAQVELVGMIFDVESMKALVEYIDLQKMLGKLSKROIQAAYSITLSEV 714  
 QY 236 EAMKNPTGQGSLELSLSCFTVTPHNEGRSRPPINSPDVLOAKKMDMLVLADIQLAQ 295  
 Db 715 QQAVSGSDSQIL-DLSNRFTYLPDHFQMKKPPLLNNADSVQAKVEMDLNLLDIEVAY 773  
 QY 296 TIQAAPGEEVEEVPHPDLRDYQLLRQQLLDGSESEYKAIOYTLKQ-----TGSYR 351  
 Db 774 SLLRGSSDDSSK-----DPIDVNYEKLKTDIKVVDRDSEAEVIRKYVKNTHATTHAYD 828  
 QY 352 CNPLRWKVNKEGDRQAQSKGNRRLLNHWGTVNAVVAAILTSGLRIMPH-----SGG 407  
 Db 829 L-EVMDIFKIERGESQRYKPKFKLHNRRLHWSRTTNFAGTSLQGLRIAPPEAVPTGY 887  
 QY 408 RVKGIYFASENKSGAGYVYTHMGCGHGVYMFGLGEVKGKHEHITIDPDLKSPGPD 467  
 Db 888 MFGKGIYFADWYKSKSANYCHTQSG-DGPIGLILLGEVALGNMYELK-HASHISKLPKGKH 944  
 QY 468 SVIARQGTPEPQAQDIELELQGVVVPQVPPVQCPQS-FKSSFSQSEYLIYKESQCLRL 526  
 Db 945 SVKGLGKTPDPSPASITLSE---GVEVPLGTGI--PSGVNDICLLYNEYIVYDIAQVNLK 998  
 QY 527 YLLEI 531  
 Db 999 YLLKL 1003

RESULT 3  
 PRO2\_MOUSE  
 ID PPO2\_MOUSE STANDARD; PRT; 559 AA.  
 AC O88554; Q99N29;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-riboyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2) (mPARP-2).  
 DE (mPARP-2).  
 GN ADPRTL2 OR PARP2 OR ADPRT2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
 RC TISSUE=Embryo;  
 RX MEDLINE=99292755; PubMed=10364231;  
 RA Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apio F., Decker P., Muller S., Hoger T., Menissier-de Murcia J., de Murcia G.M.;  
 RT "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase.";  
 RL J. Biol. Chem. 274:17860-17868 (1999).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=129Sv;  
 RX MEDLINE=21179160; PubMed=11133988;  
 RA Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M., Niedergang C.P.;  
 RT "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2 (PARP-2) gene to the gene for RNase P RNA.";  
 RL J. Biol. Chem. 276:11092-11099 (2001).  
 RN [3]

RP SEQUENCE OF 9-559 FROM N.A.  
 RC STRAIN=129/Sv X C57BL/6;  
 RX MEDLINE=99268466; PubMed=10338144;  
 RA Bergnammer H., Eder M., Marksteiner R., Auer B.;  
 RT "pADPRT-2: a novel mammalian polymerizing (ADP-ribose) transferase gene related to truncated pADPRT homologues in plants and Caenorhabditis elegans.";  
 RL FEBS Lett. 449:259-263 (1999).  
 CC -!- FUNCTION: HAS DNA-DEPENDENT POLY(ADP-RIBOSE) POLYMERASE ACTIVITY. SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribose)(N)-acceptor = nicotinamide + (ADP-D-ribose)(N+1)-acceptor.  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in testis followed by ovary.  
 CC -!- INDUCTION: By high levels of DNA-damaging agents.  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ007780; CAA07679.1; -;  
 DR EMBL: AF191547; AAK13253.1; -;  
 DR EMBL: AF072521; AAC25415.1; ALT\_INIT.  
 DR HSP: P26446; I26.  
 DR MGD: MGI:1341112; Adprt2.  
 DR InterPro: IPR001290; PARP.  
 DR InterPro: IPR004102; PARP\_reg.  
 DR Pfam: PF00644; PARP; 1.  
 DR Pfam: PF02877; PARP\_reg; 1.  
 DR Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 KW ADP-riboylation.  
 FT DNA\_BIND 1 65 POTENTIAL.  
 FT DOMAIN 66 559 NAD-BINDING (BY SIMILARITY).  
 FT DOMAIN 3 9 NUCLEAR LOCALIZATION SIGNAL 1ST PART  
 FT DOMAIN 33 39 NUCLEAR LOCALIZATION SIGNAL 2ND PART  
 FT CONFLICT 82 82 L -> V (IN REF. 2).  
 FT CONFLICT 177 177 V -> I (IN REF. 2).  
 FT CONFLICT 486 486 R -> Q (IN REF. 2).  
 FT SEQUENCE 559 AA; 63396 MW; E0AEDAE412C1445 CRC64;  
 SQ  
 Query Match 25.2%; Score 708; DB 1; Length 559;  
 Best Local Similarity 34.4%; Pred No. 51e-44;  
 Matches 200; Conservative 88; Mismatches 212; Indels 82; Gaps 23;  
 QY 1 MAPKKAS-----VQTEGSKKQKQSTG--EEDS--FRSTAEALRAAP-----ADN 41  
 Db 1 MAPRRQSGSGRRVNE-AKKVDNGNKATEDDSDPPCKMKRTQCKGPMAGGADRTKDN 59  
 QY 42 R-----VIRVDPSPFSPNPG---IQVHEDYDCTLNQTNIGNNNKFFIQLLEE 88  
 Db 60 RDSVTKLLKKGAPVDPECAAKLGAHVYCEGDVYDVNMLNQTNLQFNNNKYYLIQLLED 119  
 QY 89 GSR--FFECNWRGVGEVGVGSKMNHFTC---LEDKKDKFKKFWKTKNKEEDRDFVAQ 143  
 Db 120 DAQRNFSVWRWGRVGTGQHSLS--VTCGDLNKAKEIFQKKFLDKTKNWEEDRENFVK 177  
 QY 144 PNKYTLIEVGEAESQEAAYVKALSPQVDSGVRVTVKPCSLDPATQNLITNIFSKEMFN 203  
 Db 178 PGKYDMLQNDYAASTQD-----ESKTEETLKPSQLDLRVQELKLCINQVTMEE 229  
 QY 204 AMLMNLQDKMPLGKLTQQTARGFEALEAEAMKNPTGD-QGSLELSGCFYTVIHP 262  
 Db 230 MMENKYDTKRAPLGLKLTVAQIKAGYQSLKIEDICR--AGQHGRLVACNEAFYTRIPH 287



domain reveals amino acids involved in polymer branching.";

Biochemistry 36:12147-12154(1997).

FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.

CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose] + (N)-acceptor -> nicotinamide + [ADP-D-ribose] + (N+1)-acceptor

COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.

SUBUNIT: HOMODIMER (Potential).

SUBCELLULAR LOCATION: Nuclear.

MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN LENGTH OF 20-30 UNITS.

SIMILARITY: BELONGS TO THE PARP FAMILY.

SIMILARITY: CONTAINS 1 BRCT DOMAIN.

-----

THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; X16674; CAA34663.1; .

EMBL; M18112; AAA60137.1; .

EMBL; J03473; AAB59447.1; .

EMBL; M17081; AAA51599.1; ALT\_SEQ.

EMBL; M32721; AAA60155.1; .

EMBL; M29786; AAA51663.1; .

EMBL; M29545; AAA51663.1; JOINED.

EMBL; M29766; AAA51663.1; JOINED.

EMBL; M29767; AAA51663.1; JOINED.

EMBL; M29768; AAA51663.1; JOINED.

EMBL; M29769; AAA51663.1; JOINED.

EMBL; M29770; AAA51663.1; JOINED.

EMBL; M29771; AAA51663.1; JOINED.

EMBL; M29772; AAA51663.1; JOINED.

EMBL; M29773; AAA51663.1; JOINED.

EMBL; M29774; AAA51663.1; JOINED.

EMBL; M29775; AAA51663.1; JOINED.

EMBL; M29776; AAA51663.1; JOINED.

EMBL; M29777; AAA51663.1; JOINED.

EMBL; M29778; AAA51663.1; JOINED.

EMBL; M29779; AAA51663.1; JOINED.

EMBL; M29780; AAA51663.1; JOINED.

EMBL; M29781; AAA51663.1; JOINED.

EMBL; M29784; AAA51663.1; JOINED.

EMBL; M29785; AAA51663.1; JOINED.

EMBL; M29544; AAA51663.1; JOINED.

EMBL; M29782; AAA51663.1; JOINED.

EMBL; X56140; CAA39606.1; .

EMBL; X56141; CAA39606.1; JOINED.

EMBL; M60436; AAA60000.1; .

PIR; A26901; A26901.

PIR; A28498; A28498.

PIR; A29725; A29725.

PIR; A35635; A35635.

PIR; A33321; A33321.

PIR; B33321; B33321.

PIR; A39976; A39976.

PIR; S14010; S14010.

HSP; P26446; IA26.

Aarhus/Ghent-2DPAGE; 1620; NEPHEGE.

TM; 173870; .

InterPro; IPR001357; BRCT.

InterPro; IPR001290; PARP.

DR InterPro; IPR004102; PARP\_reg.

DR InterPro; IPR001510; Znf-PARP.

DR Pfam; PF00533; BRCT; 1.

DR Pfam; PF00644; PARP; 1.

Query Match 25.1%; Score 705; DB 1; Length 1013;

Best Local Similarity 33.8%; Pred. No. 1.9e-43;

Matches 183; Conservative 101; Mismatches 218; Indels 40; Gaps 17;

Qy 1 MAPRKASVQTEGSKKQKQKQTEEDSFSTAEALRAAPADNRVIRVDPSCFSPNRPGIOV 60

Db 492 VAPRGKSGA---ALSXKSKGVQKEEINGKSEKRMKLTLLKGAAVDPDSGLEHSHARLEKG 548

Qy 61 HEDYDCTLNQTNIGNNNKFNIIQLLEG--SRFFCNMRGVRGEV-GOSKMNHTCTLED 117

Db 549 GKVSATLGLVDIVKGNYSYKQLLEDDEKRWFRSGVRGTVIGSNKLEQMPKSKED 608

Qy 118 AKKQFKKFKTEKTKNKEERDRFVAQPNKYTLIEVQGEAESEAEAVVKAALSPVDSPGVPRT 177

Db 609 AIEHFMKLYEEKTGNAMHSKN-FTKYPKFFPLEID-YGODEEAVKKL---TVNPGTKSK 663

Qy 178 VVKPCSLDPATQNLITNIESKEMFNKMTLMNLDVKKMPLGKLTQOIARCFEALALEE 237

Db 664 LPKP-----VQDLIKMIFDVESMKKAMVEYEDLQKMPGLGKSKRQIAAAYSILSEVQQ 717

Qy 238 AMKNPTGDGQLEELSSCFYTVIPHNFRSRPPPINSDVLAQAKKMDLLVLADIQAOTL 297

Db 718 AVSGSSDSQL-DLSNRFYTLIPHDFGKMPKPLNNADSVQAKVEMLDNLLDIEVAYS 776

Qy 298 QAAPGEEEEKVEVPHLDRLDYQLLRQQLDLSGSESYKAIQTYLKO----TNSVRCRP 353

Db 777 LRGGSDSSK-----DPIDVNYEKLKTDIKVVDNRDSEAEIIRKYVKNTHATTNAYDL- 830

Qy 354 NLRHVWKNREGEDRFQAHKSKLGNRRLLWHGTWVAVVAAILTSGLRIMPH----SGGRV 409

Db 831 EVIDIFKIEREGECORYKPFKQLHNRLLWHGSRRTTFAGILSQGLRAPPAPVPTGYMF 890

Qy 410 KGIIYFASENSKSAGYVVTMHCGHQGVGMFLGEVALGKEHHITDDPSLSPGPGFDSV 469

Db 891 KGIIYFADMYSKSANYCHTSQ--GDTGLILLGEVALGNMYELK-HASHISKLPKGRHSV 947

Qy 470 IARGQTEPDPAQDIELELDGQPVVVGPPVQCPFSKSSFSQSEYLYTKESQCLRYLL 529

Db 948 KGLGKTTDPDSANI--SLDG--VDVPLGTGIS-SGVNDTSLLYEYIVYDIAQVNLKYL 1002

Qy 530 EI 531

Db 1003 KL 1004

## RESULT 5

PPOL\_BOVIN

ID PPOL\_BOVIN STANDARD; PRT; 1015 AA.

AC P18493; Q9TS00;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Poly [ADP-ribose] polymerase-1 {EC 2.4.2.30} (PARP-1) (ADPRT) (NAD(+)

DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).

GN ADPRT.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;

OC Bovidae; Bovinae; Bos.

OC [C1]

RN NCBI\_TaxID=9913;

RX MEDLINE=90382673; PubMed=2119324;

RA Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;

RT "Cloning of a full-length cDNA encoding bovine thymus

RT poly(ADP-ribose) synthetase: evolutionarily conserved segments and

RT their potential functions.";

RL Gene 90:249-254(1990).

[2]  
RN SEQUENCE OF 647-714 AND 838-903 FROM N.A.  
RP MEDLINE=88151954; PubMed=2450019;  
RA Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,  
Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;  
RT "Depression in gene expression for poly(ADP-ribose) synthetase during  
the interferon-gamma-induced activation process of murine macrophage  
tumor cells";  
RL Eur. J. Biochem. 171:571-575(1988).  
CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-ribosyl)(N)-acceptor ->  
CC nicotinamide + (ADP-D-ribosyl)(N+1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBUNIT: HOMODIMER (Potential).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: D90073; BAA14114.1; -;  
CC EMBL: X06986; CAA30046.1; -;  
CC EMBL: X06987; CAA30047.1; -;  
CC PIR: JS0428; JS0428.  
CC PIR: S00328; S00328.  
CC HSP: P26446; 1A26.  
CC InterPro: IPR001357; BRCT.  
CC InterPro: IPR001290; PARP.  
CC InterPro: IPR004102; PARP\_reg.  
CC InterPro: IPR001510; Znf-PARP.  
CC Pfam: PF00533; BRCT; 1.  
CC Pfam: PF00644; PARP; 1.  
CC Pfam: PF02877; PARP\_reg; 1.  
CC Pfam: PF00645; zf-PARP; 2.  
CC ProDom: PD004675; Znf-PARP; 2.  
CC SMART: SM00292; BRCT; 1.  
CC PROSITE: PS0172; BRCT; 1.  
CC PROSITE: PS00347; PARP\_ZN\_FINGER\_1; 2.  
CC PROSITE: PS00064; PARP\_ZN\_FINGER\_2; 2.  
CC Transferrase: Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
CC ADP-ribosylation; zinc-finger; zinc.  
CC INIT\_MET 0 0 BY SIMILARITY.  
CC  
CC DNA\_BIND 1 374 AUTOMODIFICATION DOMAIN.  
CC DOMAIN 375 525  
CC DOMAIN 386 462  
CC DOMAIN 526 1015  
CC ZN\_FING 20 55  
CC ZN\_FING 127 164  
CC DOMAIN 209 211  
CC DOMAIN 223 228  
CC MOD\_RES 408 408  
CC MOD\_RES 414 414  
CC MOD\_RES 436 436  
CC MOD\_RES 445 445  
CC MOD\_RES 446 446  
CC MOD\_RES 449 449  
CC MOD\_RES 457 457

FT MOD\_RES 472 472 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 485 485 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 489 489 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 492 492 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 514 514 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 515 515 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 521 521 ADP-RIBOSYL[N] (POTENTIAL).  
SQ SEQUENCE 1015 AA; 113355 MW; 0A5FE9D9F04F5B04 CRC64;  
  
Query Match 25.0%; Score 702; DB 1; Length 1015;  
Best Local Similarity 34.2%; Pred. No. 3.2e-43;  
Matches 187; Conservative 95; Mismatches 215; Indels 50; Gaps 20;  
  
QY 1 MAPKRKASVOTEGSKKOROGTEEDSFRSTAE--ALRAAPADNRVIRVDPSPCFSPRNG 57  
DB 494 VGPGRKSGAAP--SKSKGPVREEGTNKSEKMKLTKGAA-----VDPDSGLEHNAH 545  
QY 58 I--QVHEDYDCTLNQTNIGNNNKFFIIQLLEEG--SRFFCWNRMGRGVEV-GQSKMNH 112  
DB 546 VLEKGGKVESATLGLVDIVKGTNSYKQLLEDDEKESRYWIFRSGRVGTIGSNKLEQM 605  
QY 113 TCLEDAKDFKFKKWEKTKNKEERDRFVAQNKYTLIEVQGEAESQAVVKALESQVDS 172  
DB 606 PSKDAIEHFMKLYEEKTGNNAHKN-FTKHPKKFYPLEID-YGDEEAVKKL---TVNP 660  
QY 173 GPVRTVVKPCSLDPATONLITNIFSKEMFNAMTLMNLDVKKMPLGLTKQOIARFEAL 232  
DB 661 GYKSLPKP-----VONLIKMFVESKKAMVEIDLQKMPGLKSLKQIOAAYSIL 714  
QY 233 EALEEAMKPTDQGSLEELSCFFYTVPHNFGSRPPPINSDVLOAKKMDLLVLADIE 292  
DB 715 SEVQALSGQSSDSHTL-DLSNRFYTLPHDFGMKKPPLLNANSVQAKVEMLDLLDIE 773  
QY 293 LAQTLOAAEGEEKEVEEYPHPLDRDYQLLRCOLOLLDSESEYKAIOYTLKQ---TGN 348  
DB 774 VAYSLLRGGSDSSK-----DPIDVNYEKLTDIKVKKDSEAEIIRKYVKNTHATTHN 828  
QY 349 SYRCPNLRHVKNVREGEDRFQAHKSLGNRRLLHGTNVAVVAAILTSLGRIMPH--- 404  
DB 829 AYDL-EVVDIFKIEREGESQRYKPKQLHNRLLHSGRTTFAGILSQGLRIAPPEAPV 887  
QY 405 SGRGVKGIFYFASSENSKAGYVTTMCGGHQVGYMFLGVALGKEHHITIDDDSLKSPPP 464  
DB 888 TGYMFGKGYIFADVMVSKSANYCHTSQ--GDPIGLILLGEAALGNMYELK-HARHISKLPK 944  
QY 465 GFDSVIARTGOTPDPAQDIELELDGQPVVYVQGPVQCPFKSSFSQSESYLIYKESQCR 524  
DB 945 GKHSVKGKLGKTTDPDSASI--TVDG--VEVPLGTGIS-SGVNDTCLLYNEYIYDIAQVH 999  
QY 525 LRYLLEI 531  
DB 1000 LKYLAL 1006  
  
RESULT 6  
PPO2\_HUMAN STANDARD: PRT; 583 AA.  
ID PPO2\_HUMAN Q9Y6C8; Q9Y6C8; Q9UMR4;  
AC Q9UGN5; Q9Y6C8; Q9UMR4;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Poly (ADP-ribose) polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-  
DE ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2)  
DE (hPARP-2).  
GN ADPRT12 OR PARP2 OR ADPRT2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM 2).  
RC TISSUE=Fetal brain;



FT	VARSPLIC	68	80	MISSING (IN ISOFORM 2).
FT	CONFLICT	447		P -> H (IN REF. 2).
FT	CONFLICT	481	481	N -> H (IN REF. 4).
SEQ	SEQUENCE	583 AA;	66205 MW;	5B7AE8AE531836AF CRC64;

  

Query Match		24.8%	Score 696.5;	DB 1;	Length 583;
Best Local Similarity		33.1%;	Pred. No. 3.7e-43;		
Matches 185;	Conservative 97;	Mismatches 200;	Indels 77;	Gaps	
QY	13 GSKKQROGTEED-----SPRSTAEALRA-----	ADPNRVIRVDSPCFSRNPG-	57		
Db	54 GGRKANDRTKQDQMPGRSWASKRVSSEYKALLKGGAP-----	VDPECTAKYCKAH	106		
QY	58 --IQVHEDYDCTNOTWIGNNNKFFYIQLLLEGSR--	FFCWNRWGRVGVSGVSGKNHFT	113		
Db	107 VYCEGNDYDVMLNQTWLNQFNKKYYLIQLEDDAQRFNSVNRWGRVCKMGHSL--	VVA 164			
QY	114 C---LEDAKDFKKFWEKTKNKEERDRFVAOPNKYTLIEV-----	QGAESOEAV	162		
Db	165 CSGNLNKAKEIFQKKFLDKTKNNWEDREKPEKPGKYDMLQMDYATNTQDEETKKEESL	224			
QY	163 VKALSPQVDSGVRVTYVVKPSLDPAQTQNLITNIFSEMEFNAMTLMNLDVKKMPLGKLTK	222			
Db	225 KSPKPE-----SOLDLRVQELIKLICNVQAMEEMMKYNTTKAPLGKLTV	272			
QY	223 QQTARGFEALEALEAMKNPTGD--GQSLSELSCTVTIVPHNFSRPPPPINSDDVLOAK	281			
Db	273 AQIKAGIQSUKKIEDCIR--AGOHGRALMACNEFTYTRIPHDPLGLTPPIURTKELSEK	330			
QY	282 KDMLLVLADIATQTLQAAPGEEEEKVEEYHPHLDROYQLLRQCLQLLDSGESEYKAIO	341			
Db	331 IQLLEALGDIEAIKL-----VKTELOSPPEHPDQHYRNLCALRPLDHSYEFKVISQ	384			
QY	342 YLKQT---GNSYRCPNLRHWKNRBEGBDQFAHSLGNRRLLHWGTNVVAVAILTSG	398			
Db	385 YLQSTHAPTHSDYTTWTLDDLFVEYKDGKEKAFR--EDLHNRLMLLHWGRSMNVGILSHG	442			
QY	399 LRTMPH----SGGRVKGKGIYFASENSKSACYVTTMCHGGHQVCYMFGLGEVALGKEHHIT	454			
Db	443 LRTAPPAPITGTFMGIGIYFADMSSKSNACYFASRL--KNTGLLLSEVALGQCQCNELLE	500			
QY	455 DDPFLSKPPPPDFDSVIARGOTEPDPAQDIELELDGQPVVPGPPVQCPSFKSSSF--	SQ 512			
Db	501 ANPKAEGLQKGKSTKGLGKMAFSSAHFV--TLNGS--TVPLGPASDTGLINPDGVTLYN	556			
QY	513 SEYLIYKESOCRLRYLLEI	531			
Db	557 NEYIVNPQNVRMYLLKV	575			

  

RESULT 7		
PPOL_RAT	STANDARD;	PRT; 1013 AA.
ID	PPOL_RAT	
AC	P27008; O35937;	
DT	01-AUG-1992 (Rel. 23, Created)	
DT	15-JUL-1998 (Rel. 36, Last sequence update)	
DT	16-OCT-2001 (Rel. 40, Last annotation update)	
DE	Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)	
GN	ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).	
OS	Rattus norvegicus (Rat).	
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	
OX	NCBI_taxid=10116;	
ON	[1]	
SEQUENCE FROM N.A.		
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;	
RX	MEDLINE=98046546; PubMed=9385436;	
RA	Beneke S., Meyer R., Buerkle A.;	
RT	"Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly	
RL	(ADP-ribose) polymerase."	
RL	Biochem. Mol. Biol. Int. 43:755-761(1997).	

[2]  
RA REVISION TO 811.  
RA Beneke S., Meyer R., Buerkle A.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE OF 1-11 FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Prostate;  
RX MEDLINE=92290013; PubMed=1601134;  
RA Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,  
RA Duchaine C., Poirier G.G.;  
RT "Structural analysis of the putative regulatory region of the rat  
RT gene encoding poly(ADP-ribose) polymerase.";  
RL FEBS Lett. 302:269-273(1992).  
[4]  
RN SEQUENCE OF 514-1013 FROM N.A.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE=Prostate;  
RX MEDLINE=90027702; PubMed=2508731;  
RA Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;  
RT "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase  
RT catalytic domain and analysis of mRNA levels during the cell cycle.";  
RL Biochem. Cell Biol. 67:653-660(1989).  
CC !- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC !- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-riboseyl)(N)-acceptor =  
CC nicotinamide + (ADP-D-riboseyl)(N+1)-acceptor.  
CC !- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC !- SUBCELLULAR LOCATION: Nuclear.  
CC !- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC !- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC !- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U94340; AAC53544.1; -;  
DR EMBL; X65496; CAA46477.1; -;  
DR EMBL; X65497; CAA46478.1; ALT\_INIT.  
DR HSSP; P26446; 1A26.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; Znf-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS50172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_2; 2.  
DR TRANSFERASE; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc.  
FT INIT\_MET 0 0 BY SIMILARITY.  
FT INIT\_BIND 1 372  
FT DOMAIN 385 461 BRCT.  
FT DOMAIN 373 523 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 524 1013 NAD-BINDING.  
FT ZN\_FING 20 55 PARP-TYPE.  
FT ZN\_FING 124 161 PARP-TYPE.

FT	DOMAIN	206	208	NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT	DOMAIN	220	225	NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT	MOD_RES	1	1	BLOCKED (BY SIMILARITY).
FT	MOD_RES	407	407	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	413	413	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	435	435	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	437	437	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	444	444	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	445	445	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	456	456	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	484	484	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	488	488	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	491	491	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	512	512	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	513	513	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	519	519	ADP-RIBOSYL[N] (POTENTIAL).
FT	CONFLICT	638	638	Y -> H (IN REF. 4).
FT	CONFLICT	641	641	E -> A (IN REF. 4).
FT	CONFLICT	752	752	N -> D (IN REF. 4).
SQ	SEQUENCE	1013 AA;	112529 MW;	AA566F2B29BE97C0 CRC64;

Query Match 24.7%; Score 694.5; DB 1; Length 1013;  
Best Local Similarity 33.5%; Pred. No. 1.le-42;  
Matches 182; Conservative 101; Mismatches 216; Indels 45; Gaps 18;

QY	1	MAPKKKASVOTEGSKKQROGTEEDSFRTAEALRAAPADNRVIRVDPDSCPFSSRNFGIOV	60
DB	493	VYPKGSAPSKSK---GAVKEGVNKKSEKMKLTLAGGAADVDPDGLSESAHVLEKG	548
QY	61	HEDYDCTLTQNTIGNNNKFIQILLE--EGSRFFCWNWRGVGEV-GOSKMNHFTCLEL	117
DB	549	GVFSATGLDIVKVTGNTSYKLLLESDEKESRYWIFRSWGRVGTIGSKNLEQMPKESK	608
QY	118	AKDKKKKWEKTKKKEERDRFVAQPNKYTLIEVGEAESQEAUVK-ALSPQVDSPGVR	176
DB	609	AVEHFMKLYEETGNNAHNSKN-FTKYPKKFYPLEID-YGQDEEAVKKLAVKPTKS----	662
QY	177	TVVKPCSLDPATQNLITNIFSKEMFKNMTLMNLDVKRMPLKLTQKQIARGFEALEALE	236
DB	663	-----KLKPKVQELVGMIFDVESMKKALVEYEDLQKMPGLKLSRRQIAAYSILSEVQ	716
QY	237	EMKNPTGQGSLEELSSCFYVPHNFGSRPPPIPSDVLQAKKMDMLLVADIELAQT	296
DB	717	QAVSGSSSQIL-DLSNRYTLIPHDFGKKPPLNNTDSVQAKVEMLDNLDIEVAYS	775
QY	297	LQAAPEEEKEVEYPHLDROYQLLRCLQLLDGSESEVKALQTYLKQ-----TGSNYRC	352
DB	776	LRGGSDSSK-----DPIDVNYEKLTKDVKVVDSDSEAEVIRKYVKNTHATTHAYDL	830
QY	353	PNLRHVKNVNRREGEDRFQAHSKLGNRRLLWHGTNVAVVAAILTSLGRIMPH----SGGR	408
DB	831	-EVIDIFKIEREGESQRYKPFQOLHNRLLWHGSRRTNFAGILSOGURIAPEAPVTGYM	889
QY	409	VKGIFYASENSKSGAYVTTHMCGHQVGYMFLGEVALKEHHITIDDDPSLKSPPPGFDS	468
DB	890	FGKGYFADWWSKSNYCHTSQ--GDPILGILLGEVALGNMYELK-HASHISKLPKSKHS	946
QY	469	VIARGOTEPDPAQDIELELDGQPVVVPVQPCPS-FKSSFSQSEYLLIYKESQCLRY	527
DB	947	VKGLGKTAPDPSASI--TLDG--VEVPLGTGI--PSGVNDTCLLYNEYIYDIAQVNLKY	1000
QY	528	LLEI 531	
DB	1001	LLKL 1004	

RESULT 8  
PPOL\_CHICK  
ID PPOL\_CHICK STANDARD; PRT; 1011 AA.  
AC P26446;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)

16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
GN ADPRT.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RN SEQUENCE FROM N.A.  
RP TISSUE=Oviduct;  
RX MEDLINE=91340148; PubMed=1840535;  
RA Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;  
RT "Chicken poly(ADP-ribose) synthetase: complete deduced amino acid  
sequence and comparison with mammalian enzyme sequences.";  
RL Gene 102:157-164(1991).  
RN [2]  
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.  
RX MEDLINE=96353841; PubMed=8755499;  
RA Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;  
RT "Structure of the catalytic fragment of poly(AD-ribose) polymerase  
from chicken.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).  
RN [3]  
RN X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 659-1011, AND REVISION TO  
RP 895.  
RX MEDLINE=98191351; PubMed=9521710;  
RA Ruf A., de Murcia G.M., Schulz G.E.;  
RT "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived  
from crystal structures and homology modeling.";  
RL Biochemistry 37:3893-3900(1998).  
RN [4]  
RN X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.  
RX MEDLINE=98239716; PubMed=9571033;  
RA Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;  
RT "The mechanism of the elongation and branching reaction of poly(ADP-  
ribose) polymerase as derived from crystal structures and  
mutagenesis.";  
RL J. Mol. Biol. 278:57-65(1998).  
CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribose}(N)-acceptor =  
CC nicotinamide + {ADP-D-ribose}(N+1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBUNIT: HOMODIMER (potential).  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; X52690; CAA36917.1; .  
DR PIR; JH0581; JH0581.  
DR PDB; 2PAW; 27-MAY-98.  
DR PDB; 1PAX; 15-MAY-97.  
DR PDB; 2PAX; 27-MAY-98.  
DR PDB; 3PAX; 27-MAY-98.

DR PDB; 4PAX; 27-MAY-98.  
DR PDB; 1A26; 27-MAY-98.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; zf-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS0172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
DR PROSITE; PS00664; PARP\_ZN\_FINGER\_2; 2.  
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc; 3D-structure.  
FT DNA\_BIND 1 370  
FT DOMAIN 371 522 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 382 458 BRCT.  
FT DOMAIN 523 1011 NAD-BINDING.  
FT ZN\_FING 21 56 PARP-TYPE.  
FT ZN\_FING 125 162 PARP-TYPE.  
FT DOMAIN 207 209 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT MOD\_RES 403 403 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 404 404 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 410 410 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 411 411 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 432 432 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 434 434 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 441 441 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 442 442 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 453 453 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 454 454 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 468 468 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 481 481 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 485 485 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 488 488 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 509 509 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 510 510 ADP-RIBOSYLIN (POTENTIAL).  
FT MOD\_RES 517 517 ADP-RIBOSYLIN (POTENTIAL).  
FT CONFLICT 895 895 A -> R (IN REF. 1).  
SQ SEQUENCE 1011 AA; 113520 MW; 261AED9383139144 CRC64;  
  
Query Match 24.6%; Score 691.5; DB 1; Length 1011;  
Best Local Similarity 33.9%; Pred. No. 1.8e-42;  
Matches 185; Conservative 97; Mismatches 206; Indels 57; Gaps 20;  
  
OY 5 RKASVQTEGSKKQKRG-TTEEDSFSTRSTALRAAPADNRVIRVDPCSPSRNPQIOVHED 63  
DB 497 KPANKMSAGKVKVEQGPSKSEKKMKLTIVKGAADVDPDGL--EDSAHVFEKGGKI----- 549  
OY 64 YDCTLQNTQIGNNNKFIYIQLLEEG--SRFCWNNRWGRVGEV-GOSKNHFTCLDAAK 120  
DB 550 FSATLGLVDIVKGTNSYKYLQLEDDRESRYWVFRSGRGTGVSNGKLEQMPKSDAYE 609  
OY 121 DPKKFEWETKKNKWEERDFVQAQPNKYTLIEVOGEAESQEAQVAVKALSPQVDSGPVTVVK 180  
DB 610 HFLNLVEETKGNWSHKN-FTYKPKFYPLEID-YQDEEAVRKL---TVSAGTKSKLAK 664  
OY 181 PCSLDPATONLITNFSKEMFKNAMTLMNDVKKPLGLTKQIQIARGFEALEAEAMK 240  
DB 665 P-----IODLTKMIFDVESMKAMVFEFIDLOKPLGKLRQIQOSAYSILNEVOQAVS 718  
OY 241 NPTDGGQSEELSSCFYTVIPHNFGSRPPSPDVLQAKKMDLVLADIELAOTLOAA 300  
DB 719 DCGSRSQIL-DLSNRFYTLIPHDFGKMKPPLISNLEYIOAKVQMDLNDLIDIEVAYSLRG 777  
OY 301 PGEEEKVEEVPHPLDROYLLRCQLQLLDSGESEYKAIQTLYLQ----TGNISYRCPNLR 956  
DB 778 GNEDGDK-----DPIDINYEKLRTDIKVVDKDSSEAKIIVKYNKTHAATHNAYDL-KVY 831

SQ SEQUENCE 637 AA; 72175 MW; 527A8F464605D127 CRC64;  
 Query Match 24.5%; Score 690; DB 1; Length 637;  
 Best Local Similarity 34.7%; Pred. No. 1.2e-42;  
 Matches 190; Conservative 90; Mismatches 219; Indels 48; Gaps 18;  
 QY 6 KASVOTESKKRQGTETEEEDSFRS--TAEALRAAPADNRVIRVDPSCFPSRNPQGIQVHED 63  
 DB 112 KSSNGTDEAEDDNGPFEEKKEEKIVTATKGAAVLDQWI----PDEIKSQYHVLQRGDD 167  
 QY 64 -YDCTLNQTIGNNNKFFYLIOLLECSR--FFCWNRRGRVGEVQSGKMN-HFTCLEDAK 119  
 DB 168 VYDAILNQTNRDNNKFFVQLVLESDSKTYMYVYTWGRVGVKQSKLDGPDYSDWDRAI 227  
 QY 120 KDFKKWEFKTKNWEERDRFVAOPNKYTLIEVQGEAESOEAVVKALSPQVDSGPVRTV 179  
 DB 228 EIFTNKENDTKNTWSDRKKEFIPHPKSYTWMEDYGEKENDSPVNDIPSSS-----EV 282  
 QY 180 KP--CSLDPATQNLITNIFSKEMFKNAMTLMNLDVKMPLGKLTQOITARGFEALEALEE 237  
 DB 283 KPEQSKLDRVAKFISLICNVSMQAQHMMEIGYNAWKLPGLKISKSTISKGYEVLKRIS 342  
 QY 238 AMKNPTGGOSLEBELSSCFYTVPHNFRGRPPP--INSPDVLOAKKMDMLLVADIELA- 294  
 DB 343 VIDR--YDRTRLELSGEFYTVIPHDFGFKMSQFVDTPOKLKQKTEMVEALGEIELAT 400  
 QY 295 QTLOAAPEGEKEEVPHPLDRDYQLLRCOLQLLDSGESEYKAIQYLYKOT--GNSYR 351  
 DB 401 KLLSVDPGLQD-----PLYHYQQLNCLGTLTPVGNDSSEFSVMYMWENTHAKTHSGY 453  
 QY 352 CPNLRHWKYNREGEGRFOAHSKGLGNRRLLWHTGNVAVAAITSLGIRIMPH----SGG 407  
 DB 454 TVETAQLFRASRAVEADRFOQFSSKNRMLLWHSRLTNWAGILISQGLRIAPPEAVTGY 513  
 QY 408 RVKGIIYFASSENSKASGYVTTHMC--GGHQVGYMFLGEVALGKEHHITIDDPGLK3PPP 464  
 DB 514 MFGKGVYFADMFSKSNY----CYANTGANDGVLLLCVALGDMNELLYSDYNADNLPP 568  
 QY 465 GFDSVIARGOTEPDPAODIELDGPVVVPGPPVQCPFKSSFSQSEVLIYKESOCR 524  
 DB 569 GKUSTKGVGKTAPNPSEAOTLE-DG--VVVPLGKPV-E-RCSKGMLLYNEYIVYNVEQIK 624  
 QY 525 LRYLLEI 531  
 DB 625 MRYVIQV 631  
 RESULT 10  
 PPOL\_XENLA XENLA  
 ID PPOL\_XENLA STANDARD; PRT; 998 AA.  
 AC P31669;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)-ADP-  
 DE ribosyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).  
 DE Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;  
 OC Xenopodinae; Xenopus.  
 ON NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Ovary;  
 RA Saulier-Le Drian B.M.;  
 RL Thesis (1992), University of Rennes, France.  
 RN [2]  
 RP  
 RP SEQUENCE OF 742-876 FROM N.A.  
 RX MEDLINE=93277538; PubMed=8503897;  
 RA Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,  
 RI Miwa M.;  
 RT "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)

polymerase from *Xenopus laevis* and cherry salmon using heterologous oligonucleotide consensus sequences.";

Biophys. Res. Commun. 193:119-125(1993).

1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.

2- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose]-(N)-acceptor -> nicotinamide + {ADP-D-ribose}-(N+1)-acceptor.

3- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.

4- SUBCELLULAR LOCATION: Nuclear.

5- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES, AND BRAIN. LOW IN LIVER.

6- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN LENGTH OF 20-30 UNITS.

7- SIMILARITY: BELONGS TO THE PARP FAMILY.

8- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; Z12139; CAA78126.1; -  
 EMBL; D13810; BAA02966.1; -  
 PIR; S31735; S31735.  
 HSP; P26446; IJ26.  
 InterPro; IPR001357; BRCT.  
 InterPro; IPR001290; PARP.  
 InterPro; IPR004102; PARP\_reg.  
 InterPro; IPR001510; Znf-PARP.  
 Pfam; PF00533; BRCT; 1.  
 Pfam; PF00644; PARP; 1.  
 Pfam; PF02877; PARP\_reg; 1.  
 Pfam; PF00645; Znf-PARP; 2.  
 ProDom; PD004675; Znf-PARP; 2.  
 SMART; SM00292; BRCT; 1.  
 PROSITE; PS00172; BRCT; 1.  
 PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
 Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 ADP-ribosylation; Zinc-finger; Zinc.  
 NON\_TER 1 1  
 DNA\_BIND <1 356 AUTOMODIFICATION DOMAIN.  
 DOMAIN 357 507 BRCT.  
 DOMAIN 369 445  
 DOMAIN 508 998 NAD-BINDING.  
 ZN\_FING 8 43 PARP-TYPE.  
 ZN\_FING 111 148 PARP-TYPE.  
 DOMAIN 193 195 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
 DOMAIN 207 212 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
 MOD\_RES 391 391 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 397 397 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 419 419 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 428 428 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 429 429 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 447 447 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 454 454 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 467 467 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 471 471 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 477 477 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 495 495 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 496 496 ADP-RIBOSYL[N] (POTENTIAL).  
 MOD\_RES 503 503 ADP-RIBOSYL[N] (POTENTIAL).

FT CONFLICT 746 746 Q -> E (IN REF. 2).  
 SQ SEQUENCE 998 AA; 111126 MW; F5A25E4A3366BAE7 CRC64;

Query Match 24.4%; Score 686.5; DB 1; Length 998;  
 Best Local Similarity 34.1%; Pred. No. 4.2e-42;  
 Matches 182; Conservative 95; Mismatches 212; Indels 45; Gaps 19;

QY 10 QTESKKRQQT-BEEDSFRSTAEALRAAPADNRVIRVDPSCPPFRNPGIQVHEDYDCTL 68  
 DB 488 KSSGKVKKEGSKNKKMLTVKGAADPDSD---ELEDSCVLETGG---KIFSATL 540  
 QY 69 NQTNIGNNNNFYIIQLLE--EGSRFFCWNRMGRVGEV-GOSKMNHFTCLEDAKDKFKK 125  
 DB 541 GLVDITRGNTYYKLQLEHDDRSRYWFRSGRWGVTVIGSKKLEEMSKEDAIEHFNL 600  
 QY 126 FWEKTKNKEEDRFVAQPNKYTLIEVQGEAESEAVYKALSPQVDSGPRVTVVVKPSLD 185  
 DB 601 YQDKTGNWHS-PNFTKYPKKFPYPLEI--DYQOEEDVVKLS--VGACTKSKLAKP--- 651  
 QY 186 PATQNLITNIFSKEMFNKNTMLMDVKMPLGLTKTKQIARGFPALAEAMKNPTGD 245  
 DB 652 --VQELIKLIFDVESMKKAMVFEIDLOKPLGLSKRKIQSAYSILSQVQQAQVSELS 709  
 QY 246 GQSLLELSSCFTVTPHNFGRSRPPINSPDVLOAKDKMLVLADIELAOTLOAAPGEE 305  
 DB 710 ARLL-DLSNQYTLPHDFGKKPPLNNLEYIAQVKOMLNDLIDIEVAYSLRGGADGG 768  
 QY 306 EKVEVPHPLDRDYQLRCQLQDLSGESEYKATQTYLK----QTGNSYRCNLRHWKV 361  
 DB 769 EK-----DPIDVKEIKTDIKVVAKDEESRIICDYVKNTHADTHNAYDLEVL-EIFK 822  
 QY 362 NREGGDFQAHSKLGNRLNHLHGTNNVAVAAITSLGLRIMPH----SGRGVKGIIYAS 417  
 DB 823 DREGYQRYKPFKQLHNNHGSRTTNFAGILSQGLRIAPPEAPVTVGYMGKGIYFAD 882  
 QY 418 ENSKSAGYVTTMHGCGHOVGYMFLGEVALGKEHHTTIDPSLKSPPPGDFSVIARGQEP 477  
 DB 883 MVSKSANYCHAM--PGSPIGLILGEVALGNHNLKAASQITKL-PKGKHSVKGIGRTAP 939  
 QY 478 DPAQDIELELDGQPVVPPQGPVQCPSPKSSFSOSEYLIYKESQCRURYLEI 531  
 DB 940 DPSATV--QLDG--VDVPLGKGTSA-NISDTSLLYNEYIVYDIAQVNLKYLKL 988

RESULT 11  
 PPOL\_MOUSE  
 ID PPOL\_MOUSE STANDARD; PRT; 1012 AA.  
 AC P11103; Q9JLX4; Q9QVQ3;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
 DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (msPARP).  
 GN ADPRT OR ADPRT1 OR ADPRT.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RP STRAIN=BXSB;  
 RX MEDLINE=89263780; PubMed=2498841;  
 RA Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;  
 RT "Sequence and organization of the mouse poly (ADP-ribose) polymerase  
 RT gene.";  
 RL Nucleic Acids Res. 17:3387-3401(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.  
 RC STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;  
 RX MEDLINE=20270268; PubMed=10809783;  
 RA Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;  
 RT "Characterization of sPARP-1. An alternative product of PARP-1 gene

with poly(ADP-ribose) polymerase activity independent of DNA strand breaks.";  
 J. Biol. Chem. 275:15504-15511(2000).  
 [3]  
 MEDLINE=96007847; PubMed=7578427;  
 Auer B., Flick K., Wang Z.Q., Haldacher D., Jaeger S., Berghammer H., Kofler B., Schweiger M., Wagner E.F.;  
 "On the biological role of the nuclear polymerizing NAD<sup>+</sup>: protein(ADP-ribose) transferase (ADPRT): ADPRT from Dictyostelium discoideum and inactivation of the ADPRT gene in the mouse.";  
 Biochimie 77:444-449(1995).  
 CC -1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + (ADP-D-riboseyl)(N)-acceptor = nicotinamide + (ADP-D-riboseyl)(N+1)-acceptor.  
 CC -1- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -1- SUBUNIT: HOMODIMER (Potential).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a short form/SPARP-1; may be produced by alternative initiation.  
 CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

-----  
 This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----

EMBL; X14206; CAA32421.1; -  
 DR EMBL; AF126717; AAF61293.1; ALT\_INIT.  
 DR PIR; S04200; S04200.  
 DR HSSP; P26446; 1A26.  
 DR MGD; MGI:1340806; Adprt1.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR001510; Znf-PARP.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF00645; Znf-PARP; 2.  
 DR ProDom; PD004675; Znf-PARP; 2.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS00172; BRCT; 1.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_2; 2.  
 DR PROSITE; PS50064; PARP\_ZN\_FINGER\_1; 2.  
 DR TRANSFERASE; Glycosyltransferase; NAD; DNA-binding; Nuclear protein; ADP-riboseylation; Zinc-finger; Zinc; Alternative initiation.  
 KW TRANSFERASE; Glycosyltransferase; NAD; DNA-binding; Nuclear protein; ADP-riboseylation; Zinc-finger; Zinc; Alternative initiation.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT CHAIN 1 1012 POLY [ADP-RIBOSE] POLYMERASE-1, LONG ISOFORM.  
 FT FT  
 FT CHAIN 521 1012 POLY [ADP-RIBOSE] POLYMERASE-1, SHORT ISOFORM.  
 FT FT  
 FT INIT\_MET 521 521 FOR SHORT ISOFORM.  
 FT DNA\_BIND 1 371 AUTOMODIFICATION DOMAIN.  
 FT DOMAIN 372 522 BRCT.  
 FT DOMAIN 384 460 NAD-BINDING.  
 FT DOMAIN 523 1012 PARP-TYPE.  
 FT ZN\_FING 20 55  
 FT ZN\_FING 124 161

FT	DOMAIN	206	208	NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT	DOMAIN	220	225	NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT	MOD_RES	406	406	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	412	412	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	434	434	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	436	436	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	443	443	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	444	444	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	447	447	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	455	455	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	483	483	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	487	487	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	490	490	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	511	511	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	512	512	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	518	518	ADP-RIBOSYL[N] (POTENTIAL).
FT	CONFLICT	590	590	L -> V (IN REF. 2).
FT	CONFLICT	607	607	E -> D (IN REF. 2).
FT	CONFLICT	611	611	Q -> H (IN REF. 2).
FT	CONFLICT	628	628	N -> D (IN REF. 2).
FT	CONFLICT	678	678	D -> E (IN REF. 2).
FT	CONFLICT	702	702	R -> F (IN REF. 3).
FT	CONFLICT	716	716	Q -> E (IN REF. 2).
FT	CONFLICT	757	757	Q -> L (IN REF. 2).
FT	CONFLICT	856	856	R -> F (IN REF. 3).
FT	CONFLICT	981	981	A -> C (IN REF. 2).
SEQ	SEQUENCE	1012 AA;	112968 MW;	4354C3E5F01B9439 CRC64;

Query Match 24.3%; Score 683.5; DB 1; Length 1012;

Best Local Similarity 32.4%; Pred. No. 7e-42;  
 Matches 177; Conservative 105; Mismatches 215; Indels 49; Gaps 17;

QY	1	MAPKRASVQTEGSKQROQTEEDSFRPAEALRAAPADNRVIRVDPSPCFSRNPGIQV	60
DB	492	VAPRGSAAPSKSK-----GCFKEEGVKNSEKRMKLTGKGAADVPDGLSHSHVLEKG	547
QY	61	HEYDCTLTNIGNNNNKFIYIIQLLEG--SRFFCWNRMGRVGEV-GOSKMNNHTCTLED	117
DB	548	GKVFSAITGLVDIVKGTNSYKQLLEDKESYWIYFRSGRLGTIVGSKNLEQMSKEE	607
QY	118	AKDFFKKFWEKYNKWEEDRFVAQPNKTLTLEVOQEAESQEAQVAVKALSPQVDSGPVRT	177
DB	608	AVEQFMKLYEKTGNAMWSKN-FTKYPKFYPLEID-YGQDEEAVKK-----L	653
QY	178	VVKP---CSLDPATONLITNIESEKMFKNAMTLMNLDVKKMPLGKLTQQTARGFEALEA	234
DB	654	TVAPGTSKLPKPVQELVGMIFDVSMMKALVEYEDLQKMPGLKLSRRQIAAYSILSE	713
QY	235	LEAMKNPTGDGGSLEELSCFYTVIPHNGRGRPPPIPSDVLQAKKMDLLVLADIELA	294
DB	714	VQPVSGSSESQIL-DLSNRFTLIPHDFGMRKPPLLNNADSVOAKVEMLDNLLDIEVA	772
QY	295	QTLQAPGEEEEKVEEVPHPDLDRDYQLLRQQLLDSESESEYKAIOYLYKQ---TCNSY	350
DB	773	YSLRGSGDDSSK-----DPIDVNYEKLTKTDIKVDRDSEAEVIRKIVYKTHATTNNAY	827
QY	351	RCPNLRHWKVNREGEDRFQAHSKLGNNRLLMHTGTVAVVAAILTSGLRIMPH---SG	406
DB	828	DL-EVIDIFKIERGESQRYKPPFQLNRRLLMHSRTTNFAGILSOGRLTAPEAPVTG	886
QY	407	GRVGKGYFASENSKSAGYVYTHHCGGHQGVYMFELGVALGKEHHITDDPSLSPPPGF	466
DB	887	YMFEGKGYFADVMYSKSNYCHTSQ--GDPICLIMGLGVALGNMYELK-HASHISKLPKKG	943
QY	467	DSVTARGCTEPDPAQDIELEDGQPVVPOGPPVQCPS-FKSSFSFOSQSELYKESQRL	525
DB	944	HSVKGGLKTPDPSASITL-----GVEVPLGTGI--PSGVNDTALLYNEYIVDIAQVNL	997
QY	526	RYLLEI 531	
DB	998	KYLLKL 1003	

RESULT 12  
PPOL\_SARPE STANDARD; PRT; 996 AA.  
AC Q11208;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
DE ribosyltransferase) (Poly[ADP-ribose] synthetase).  
OS Sarcophaga peregrina (flesh fly) (Boettcherisca peregrina).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestridea; Sarcophagidae; Sarcophaga.  
OX NCBI\_TaxID=7386;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=94170813; PubMed=8125121;  
RA Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,  
RA Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;  
RT "Cloning and functional expression of poly(ADP-ribose) polymerase  
RT cDNA from Sarcophaga peregrina".  
RL Eur. J. Biochem. 220:607-614(1994).  
CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =  
CC nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.  
CC -!- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; D16482; BAA03943.1; -  
DR HSPF; P26446; I426.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; zif-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS00172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; FALSE\_NEG.  
DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc.  
FT DNA\_BIND 1 369 BY SIMILARITY.  
FT DOMAIN 370 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 382 456 BRCT.  
FT DOMAIN 508 996 NAD-BINDING.  
FT ZN\_FING 19 54 BY SIMILARITY.  
FT ZN\_FING 126 164 BY SIMILARITY.  
FT DOMAIN 211 214 NUCLEAR LOCALIZATION SIGNAL 1ST PART.

FT DOMAIN 232 235 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
SQ SEQUENCE 996 AA; 113018 MW; 690DD36E7487298 CRC64;  
  
Query Match 22.9%; Score 645.5; DB 1; Length 996;  
Best Local Similarity 32.8%; Pred. No. 4e-39;  
Matches 175; Conservative 101; Mismatches 205; Indels 53; Gaps 18;  
  
QY 22 EEDSFSTAEALRAAPADNRV-----LRVDPSCPF-----SRNPGIQVHEDYDCTL 68  
Db 489 EESKSSKSIYTKSVKSMWTKIKDGLAVDPDSGLDVAHVYVSRN-----KEKYNVL 543  
QY 69 NQTIGNNNKFFYIQLLEEG--SRFCWNWRVG-EVGQSKMNHFTCLDEAKDKFKKK 125  
Db 544 GITDIQKNKSNFYKLQLESMDKNRFVWSGRIGTITGCKLNDLSNLVDAIVQFKEL 603  
QY 126 FWEKTKNKEERDRFVAQPNKYTLIEVOGEAESEAVVVKALSPQVDSGPVTVVVKPCSLD 185  
Db 604 YLEKSGNHENFENFVKVGRMYPIDIDYAE-----KIDLSAEHDIKSKLPL- 652  
QY 186 PATONLTINIFSKEMFNAMTLMNLDVKKMPLGLTKQOIARGFEALFEALFEAMNPTGD 245  
Db 653 -SVQDIILKMFVDSMKRTMFEFDLDMKPLGKLSQKIOSAYKVLTEIYELIQG-GGT 710  
QY 246 GQSLSELSSCFYTVPHNFGSRPPINSPDVLAQAKDMLVLADIELAQTLQAAPGEE 305  
Db 711 NAKIDATNRFTLPHNFGTQSPPLDITTEQVEQLRQMLDSLIEICAYSLLQT---ED 767  
QY 306 EKVEVPHPLDRDQOLLRCQQLDSESEYKATOTYKQT-GNSYRCPNLR--HWVKVN 362  
Db 768 SKAD--INPIDKHVEQLTKLEPLDKNSEEYILLQKYVKNTHAETHKLYLDEVVDIFKA 825  
QY 363 REGGEDRFOAHSKLGNRLLWHGTNVAVVAAILTSGLRIMPH---SGGRVKGIIYFASE 418  
Db 826 RQGEARYKPKPKLHNRKLLWHGSRNTNFAGILSHGLKIAPPEAPVTGYMGKGIYFADM 885  
QY 419 NSKAGAVYVTHMCGHGVYMFGLGEALGKEHHITIDDPKSLKSPPPGSDSVIARGQTEPD 478  
Db 886 VSKSANYCCTSH--HNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK-HSCFGRGRTMPN 942  
QY 479 PAQDIELELDQPVVVGPPVQCPSPKSSFSQSEYLIYKESOCRLRYLLEIH 532  
Db 943 PSEIIRE-DG--VEIPLGKPTINDSLK-SSLLYNEFTIYDIAQVNIQYMLRMN 992  
  
RESULT 13  
PPOL\_DROME STANDARD; PRT; 994 AA.  
AC P35875; O9W5Q5; O9W5S1;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
DE ribosyltransferase) (Poly[ADP-ribose] synthetase).  
GN PARP OR CG17696/CG17718.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93234521; PubMed=8475096;  
RA Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,  
RA Sugimura T., Miwa M.;  
RT "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:  
RT leucine zipper in the auto-modification domain".  
RL Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).  
RN [2]  
RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND  
RP TISSUE SPECIFICITY.  
RC STRAIN=CANTON-S;  
RX MEDLINE=98234380; PubMed=9565614;



RA Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;  
RT "Genomic organization of Drosophila poly(ADP-ribose) polymerase and  
RT distribution of its mRNA during development.";  
RL J. Biol. Chem. 273:11881-11886(1998).  
RN [3]

RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Branton R.C., Rogers Y.-H.C., Blazee R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,  
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirska R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Welstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).

CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =  
CC nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.  
CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES  
CC OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM  
CC IN LATER EMBRYOS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN  
CC EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.  
CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----

DR EMBL; D13806; BAA02964.1; -  
DR EMBL; AF051548; AAC24518.1; -  
DR EMBL; AF051544; AAC24518.1; JOINED.  
DR EMBL; AF051545; AAC24518.1; JOINED.  
DR EMBL; AF051546; AAC24518.1; JOINED.  
DR EMBL; AF051547; AAC24518.1; JOINED.  
DR EMBL; AE002935; AAF45400.1; -  
DR EMBL; AE002666; -; NOT\_ANNOTATED\_CDS.  
DR EMBL; AE002892; AAF45445.2; ALT\_SEQ.  
DR PIR; A47474; A47474.  
DR HSP; P26446; I26.  
DR FlyBase; FBgn010247; Parp.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; ZnF-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; zf-PARP; 2.  
DR ProDom; PD004675; ZnF-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS00172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 1.  
DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
DR Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; zinc; Alternative splicing.  
FT DNA\_BIND 1 367  
FT DOMAIN 368 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 380 454 BRCT.  
FT DOMAIN 508 994 NAD-BINDING.  
FT ZN\_FING 19 54 PARP-TYPE.  
FT ZN\_FING 123 161 PARP-TYPE.  
FT DOMAIN 208 210 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 223 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT VARSPIC 376 564 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 994 AA; 113791 MW; ACAB5A270DD29E08 CRC64;

Query Match 22.0%; Score 619; DB 1; Length 994;  
Best Local Similarity 31.5%; Pred No. 3 4e-37;  
Matches 168; Conservative 100; Mismatches 222; Indels 44; Gaps 17;

QY 16 KQCGTEEDSFRTAEALRAAPADNRVIRVDPSCPSFRNPQI-----VHED----YDCT 67  
DB 481 KSRIPKRTTKSLNSNIYTKMPV-SRTFKVKDGLAVDPDGLDIAHVVDSSNNKYSVV 539  
QY 68 LQQTNGNNNNFYIQLL--EGRSFFCWNWRGVG-EVGQSKMHHFTCLEDAKDFK 124  
DB 540 LGLTQIRKNKSNYYKQLLKADKDKYKWIFRSGWRTGTGNIGSKLEEFDTSESARKNKE 599  
QY 125 KWEKTKNWEEDRDVVAQPNKYTLLEVOGEAESQEAVALKSPQVDSGPVRTVWPCSL 184  
DB 600 IYADTKNGEYQDNFVKRTGRMYPDIEQ--YDDOKLVKHSHFTS-----KL 647  
QY 185 DPATQNLTNIFSKEMFKNMTLMNLDVKKMPLGKLTQKQIARGFEALFALEAMKNPTG 244  
DB 648 ELISVQNLKILFIDDSMNKTLMEFHIDMDKMPGLKLSAHQISQYRVVKEIYNVLECGSN 707  
QY 245 DGOSLEELSSCYTVPHPNFGSRPPINSPDVLOAKKMLVLADIELAQTLQAAPGEE 304  
DB 708 TAK-LIDATNREYTLIPHPNFGVQLPTLIETHQOIEDRLQMLDSLAEIYAYSI-----IK 761  
QY 305 EKVVEVPHPLDRDYQLLRCOLLDSDGESEYKAIQTYLKQT--GNSYRCPNLR--HWKV 361  
DB 762 SEDVDACNPLDNHYAQIKTLQVALDKNSEEFSLSQYKVNTHASTHKSYDLKIVDFKV 821



```
QY 362 NREGGRFQAHKSLGNRRLLHGTNNVAVVAAILTSLGRI-----MPHSGGRVKGIIYFAS 417
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 822 SROGEARRFKPKLHNKLLHWSRLTNFVGLSHGLRIAPPEAPTGYMFGKGIYFAD 881
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 418 ENSKAGVVTMTMCGGHQGVYMFGEVALGKEHITITDDPSLKSPPPGFDVSIARGQTEP 477
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 882 MVKSANYCTCSQ--QNSTGLMLLSEVALGDMMECT-SAKYINKLSNKKHSCFGGRTMP 938
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 478 DPAQDIELELDQVVPVPPQCPGKSSFSQSEYLIYKESQCLRYLLEI 531
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 939 DPTKSY-IRSDG--VEIPYGETIDEHLK-SSLLYNEYIYDVQAQNIQYLFM 988
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 14
YQ4 CAEEL
ID YQ4 CAEEL STANDARD; PRT; 538 AA.
AC Q09525;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-FEB-1996 (Rel. 33, Last annotation update)
DE Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.
GN E02H1.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Smith A.;
RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-
CC RIBOSYLTRANSFERASE (EC 2.4.2.30).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z47075; CAA87379.1; -.
CC HSSP; P26446; 1A26.
CC WormPep; E02H1.4; CE01539.
CC InterPro; IPR001290; PARP.
CC InterPro; IPR004102; PARP_reg.
CC Pfam; PF00644; PARP; 1.
CC Pfam; PF02877; PARP_reg; 1.
CC KW Hypothetical protein.
CC SEQUENCE 538 AA; 61268 MW; 3144E25465FC7341 CRC64;
-----
Query Match 14.1%; Score 396.5; DB 1; Length 538;
Best Local Similarity 26.1%; Pred. No. 2.1e-21;
Matches 149; Conservative 87; Mismatches 184; Indels 151; Gaps 24;

QY 57 GIVHEDYDCLNQTNGNNNNKFIYIQLLEGSRRFCWNRWGRVGEVQSGKMHFTCLE 116
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 11 GYKVH-----LCKNTNAQNNKFPYDMELIDEGDFTVKLINGRIGYRGVYTLQKDFDLD 64
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 117 DAKDKFKKFWKTKWEERDFVAQPNKYTLIEVQGEAESQEAIVKALSPQ----VDS 172
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 65 RAKKFFSKFYKTHLWEERDD-EPVNPKYAVVELATNARQTEKVEKKEPEPKVDE 123
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 173 GPV-----RTWVKPSLDPAQNLITNIFSKEMFKNAMTL-----MNLDVKKM----- 215
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 124 KNTGRKKRGIVREKKEIKEEPEVEV--NEKLKELMKICDEVDHLGLLQKLFNEAF 181
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 216 --PLGKLTQKIARGFALAEAMKNPT-----GDGQSLEELSS 254
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 182 GRPTDCLSLAQLTTGYEILSKIEESIGGKSARRSTRGRPRVADRVLAVKSDGPSLHDINK 241
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
QY 255 CFYTVIPHNFRSRPPPIPDVLQAKOMLLVL-----ADIELAQTLOAAPGEEEKVE 309
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 242 -YYSLSHSFGFCVPPPKIDSHAKIOARELLDALKGSTSEASLEKDLKKTASSKD----- 295
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 310 EVPHPLDRDQQLRCQLQLDLSGESEYKATQYTLKQTSNYSYRCPNLR---HWKVN----- 362
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 296 ----IYORLYERLPCHLEPVSE-----EIAKGIDGCLAMRGPTHYCYKLSLIDA 339
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 363 -----REGGDRFQAHKSLG-----NRRLLHGTNNVAVVAAILTS 397
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 340 FELKDPNEIPTAEVPEVQEPKPKGRKSTKTAAPTVPPTTKRLLHWGTRVTVFSILMN 399
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 398 GLR--IMPHSGGRVKGIIYFASENSKSAGVYVTTMHC--GGHQGVYMFGEVALGK----- 448
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 400 GLQFPVGDRCGLMEFGVGFANVPVTKSANVC---CPEASKRV-FMLLCEVETANPLVLY 454
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 449 EHHITIDPSLSKPPPGFDVSIARGQTEPDPAQDIELELDQVVPVPPQCPGSPKSKS- 507
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 455 ESEIDADEKMEKAKK---TSVYAAAGKHTPRDT-----VEING-----IPAFKSN 495
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
QY 508 -----SFSQSEYLIYKESQCLRYLLEI 531
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
Db 496 LETIEETRLLYDEYVMFNKEHFKIKYVEV 526
   :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||

RESULT 15
PROV_HUMAN
ID PROV_HUMAN STANDARD; PRT; 1724 AA.
AC Q9UKK3; Q75903; Q9HIM6; Q14682;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Vault poly(ADP-ribose) polymerase (EC 2.4.2.30) (VPARP) (193-kDa vault
DE protein) (PARP-related/falipai-related H5/proline-rich) (PH5P).
DE ADPRTLI OR PARLP OR KIAA0177.
GN Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 306-319.
RX MEDLINE=99408776; PubMed=10477748;
RA Kickhoefer V.A., Silva A.C., Kedersha N.L., Inman E.M., Ruland C.,
RA Streuli M., Rome L.H.;
RT "The 193 kDa vault protein, VPARP, is a novel poly(ADP-ribose)
RT polymerase."
RL J. Cell Biol. 146:917-928(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymus;
RX MEDLINE=20112770; PubMed=10644454;
RA Still I.H., Vince P., Cowell J.K.;
RT "Identification of a novel gene (ADPRTLI) encoding a potential
RT poly(ADP-ribosyl)transferase protein."
RL Genomics 62:533-536(1999).
RN [3]
RP SEQUENCE FROM N.A.
RA Tromans A.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE OF 94-1724 FROM N.A.
RX TISSUE=Bone marrow;
RC MEDLINE=96281124; PubMed=8724849;
RA Nagase T., Seki N., Ishikawa K.-I., Tanaka A., Nomura N.;
RT "Prediction of the coding sequences of unidentified human genes. V.
RT The coding sequences of 40 new genes (KIAA0161-KIAA0200) deduced by
RT analysis of cDNA clones from human cell line KG-1."
RL DNA Res. 3:17-24(1996).
RN [5]
RP DISCUSSION OF SEQUENCE.
RX MEDLINE=99198702; PubMed=10100603;
RA Jean L., Risler J.-L., Nagase T., Coulouarn C., Nomura N.,
RA Salier J.-P.;
```



---

**THIS PAGE BLANK (USPTO)**

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:01:34 ; Search time 117.38 seconds  
(without alignments)  
785.537 Million cell updates/sec

Title: US-09-701-586b-8  
Perfect score: 2813  
Sequence: 1 MAPKKASVQTEGSKKQROG.....EVLKYESQCLRYLLEIHL 533

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_19:\*
- 1: sp\_archaea:\*
  - 2: sp\_bacteria:\*
  - 3: sp\_fungi:\*
  - 4: sp\_human:\*
  - 5: sp\_invertebrate:\*
  - 6: sp\_mammal:\*
  - 7: sp\_mhc:\*
  - 8: sp\_organelle:\*
  - 9: sp\_phase:\*
  - 10: sp\_plant:\*
  - 11: sp\_rodent:\*
  - 12: sp\_virus:\*
  - 13: sp\_vertebrate:\*
  - 14: sp\_unclassified:\*
  - 15: sp\_rvirus:\*
  - 16: sp\_bacteriap:\*
  - 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2772.5	98.6	528	11 Q91YR6	Q91YR6 mus musculus
2	2241	79.7	533	4 Q96CG2	Q96CG2 homo sapien
3	822.5	29.2	612	5 Q9TX06	Q9TX06 dictyosteli
4	717	25.5	653	10 Q50017	Q50017 zea mays (m
5	700.5	24.9	1014	11 Q921K2	Q921K2 mus musculus
6	686.5	24.4	607	13 Q9PS82	Q9PS82 gallus gall
7	686.5	24.4	607	13 Q9PS81	Q9PS81 xenopus. na
8	683.5	24.3	635	10 Q81294	Q81294 arabidopsis
9	621.5	22.1	983	10 Q92P54	Q92P54 arabidopsis
10	621.5	22.1	1009	10 Q9SJM4	Q9SJM4 arabidopsis
11	610	21.7	593	5 Q9TX05	Q9TX05 drosophila
12	584.5	20.8	969	10 Q24570	Q24570 zea mays (m
13	584.5	20.8	980	10 Q92SV1	Q92SV1 zea mays (m
14	569.5	20.2	945	5 Q9N4H4	Q9N4H4 caenorhabdi
15	475.5	16.9	727	5 Q9XUA5	Q9XUA5 caenorhabdi
16	310.5	11.0	2276	5 Q9TXQ1	Q9TXQ1 caenorhabdi

17	275.5	9.8	815	10 Q9SWB4	Q9SWB4 glycine max
18	263.5	9.4	815	10 Q9FK91	Q9FK91 arabidopsis
19	156.5	5.6	1181	5 Q9XZ37	Q9XZ37 drosophila
20	156.5	5.6	1181	5 Q9VBF3	Q9VBF3 drosophila
21	146	5.2	1327	4 Q95271	Q95271 homo sapien
22	127.5	4.5	363	4 Q9H8R9	Q9H8R9 homo sapien
23	125.5	4.5	5198	5 Q76518	Q76518 caenorhabdi
24	124.5	4.4	1166	4 Q9H2K2	Q9H2K2 homo sapien
25	124.5	4.4	1265	4 Q9HAS4	Q9HAS4 homo sapien
26	120.5	4.3	181	12 Q55721	Q55721 chilo iride
27	119.5	4.2	954	10 P93826	P93826 arabidopsis
28	118.5	4.2	1092	2 Q50236	Q50236 zymomonas m
29	115	4.1	451	5 Q9V635	Q9V635 drosophila
30	114.5	4.1	1203	16 Q9CJ19	Q9CJ19 lactococcus
31	112.5	4.0	935	10 Q9CA26	Q9CA26 arabidopsis
32	111.5	4.0	261	4 Q9H8F2	Q9H8F2 homo sapien
33	111.5	4.0	359	4 Q9Y4P7	Q9Y4P7 homo sapien
34	111.5	4.0	1763	11 Q9JXK5	Q9JXK5 mus musculu
35	110	3.9	924	4 Q969W4	Q969W4 homo sapien
36	110	3.9	1223	4 Q9UFT5	Q9UFT5 homo sapien
37	108.5	3.9	1065	2 Q9AHK8	Q9AHK8 borrelia bu
38	108	3.8	949	16 Q92HW8	Q92HW8 rickettsia
39	108	3.8	1342	10 Q9FKN5	Q9FKN5 arabidopsis
40	107.5	3.8	920	5 Q9U1M9	Q9U1M9 dictyosteli
41	107	3.8	1327	11 Q61595	Q61595 mus musculu
42	106.5	3.8	1412	4 Q96RT1	Q96RT1 homo sapien
43	106.5	3.8	1435	3 Q03291	Q03291 saccharomyc
44	106	3.8	757	16 Q9A0Q1	Q9A0Q1 streptococc
45	106	3.8	769	10 Q9C6G1	Q9C6G1 arabidopsis

ALIGNMENTS

RESULT 1

Q91YR6 PRELIMINARY; PRT; 528 AA.  
AC Q91YR6;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 59.4 KDA PROTEIN.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014870; AAL14870.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;

Query Match 98.6%; Score 2772.5; DB 11; Length 528;  
Best Local Similarity 98.9%; Pred. No. 1.3e-216;  
Matches 527; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

Qy	1	MAPKKASVQTEGSKKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSSRNPGIQV	60
Db	1	MAPKKASVQTEGSKKQROGTEEDSFRSTAEALRAAPADNRVIRVDPSCPFSSRNPGIQV	60
Qy	61	HEDYDCTLNQTNIGNNNKFFYIIQLLEGSRRFCFNRNMGVRGVEGQSKMNHFTCLD	120
Db	61	HEDYDCTLNQTNIGNNNKFFYIIQLLEGSRRFCFNRNMGVRGVEGQSKMNHFTCLD	120
Qy	121	DFKKKFEKTKNKNWEERDFVAQPNKYTLIEVQGAESQAEVVKALSPQVDSGPRTVVVK	180
Db	121	DFKKKFEKTKNKNWEERDFVAQPNKYTLIEVQGAESQAEVVKALSPQVDSGPRTVVVK	175
Qy	181	PCSLDPATONLITNFFSKEMFKNAMTLMNLDVKKMPLGLTKQQTARGFEALAELEAMK	240
Db	181	PCSLDPATONLITNFFSKEMFKNAMTLMNLDVKKMPLGLTKQQTARGFEALAELEAMK	240

```
Db 176 PCSLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQKIARGFEALEEAMK 235
QY 241 NPTGDGSLSELSCTFTVTPHNFGRSRPPINSPDVLQAKKMDLLADIELAQTLQAA 300
Db 236 NPTGDGSLSELSCTFTVTPHNFGRSRPPINSPDVLQAKKMDLLADIELVQTLQAA 295
QY 301 PGEERKEVEEVPPLDRDYQLLRQQLQDLSGESEYKAIQTYLKQTSYRCPNLRHWK 360
Db 296 PGEERKEVEEVPPLDRDYQLLRQQLQDLSGESEYKAIQTYLKQTSYRCPNLRHWK 355
QY 361 VNRGEGDRFOAHKSLGNRRLLWHTGNVAVVAAILTSLGRIMPHSGRGVKGIFYASNS 420
Db 356 VNRGEGDRFOAHKSLGNRRLLWHTGNVAVVAAILTSLGRIMPHSGRGVKGIFYASNS 415
QY 421 KSAGYVTMTGCGHGVGMFLGEVALGKEHHITIDDPSSLKSPPPGFSVARGOTEPDPA 480
Db 416 KSAGYVTMTGCGHGVGMFLGEVALGKEHHITIDDPSSLKSPPPGFSVARGOTEPDPA 475
QY 481 QDIELELDGQPVVPPQPCPSFKSSFSQSEYLYKESQCRRLYLLEIHL 533
Db 476 QDIELELDGQPVVPPQPCPSFKSSFSQSEYLYKESQCRRLYLLEIHL 528

RESULT 2
Q96CG2 ID Q96CG2 PRELIMINARY; PRT; 533 AA.
AC Q96CG2:
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE HYPOTHETICAL 60.1 KDA PROTEIN.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014260; AH14260.1; -.
KW Hypothetical protein.
SQ SEQUENCE 533 AA; 60089 MW; 6296A0E439CC7767 CRC64;

Query Match 79.7%; Score 2241; DB 4; Length 533;
Best Local Similarity 80.1%; Pred. No. 1.9e-173;
Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

QY 1 MAPKKASVQTEG--SKKQKQTEEDSFRSTAEALRAAPADNRVIRVDPSCPSFRNPGI 58
Db 1 MAPKKPWWQTEGPEKKKGRQAGREDDPFRSTAEALKAIPAERIRVDPCTPLSSNPGT 60

QY 59 QVHEDYDCTLNQTNIGNNNKFIYIQLLEGRSFF-CWNRWGRVGEVQSGKMHFTCLEL 117
Db 61 QVHEDYDCTLNQTNIGNNNKFIYIQLLEGRSFF-CWNRWGRVGEVQSGKMHFTCLEL 120

QY 118 AKKDFKKFKWEKTKNWEERDRFVAQPNKYTLLEVOGEAESQAVKALSPQVDSGPVRT 177
Db 121 AKKDFKKFKWEKTKNWEERDRFVAQPNKYTLLEVOGEAESQAVKALSPQVDSGPVRT 175

QY 178 V---VKPCSLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQKIARGFEALEA 234
Db 176 VTRKVPQCSLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLSKQIARGFEALEA 235

QY 235 LEAMKNPTGDGQSLSELSCTFTVTPHNFGRSRPPINSPDVLQAKKMDLLADIELA 294
Db 236 LEALKQPTGDGQSLSELSCTFTVTPHNFGRSRPPINSPDVLQAKKMDLLADIELA 295

QY 295 QTLQAAPEEERKEVEEVPPLDRDYQLLRQQLQDLSGESEYKAIQTYLKQTSYRCPN 354
Db 296 QALQAV-SEGEKTVVEEVPPLDRDYQLLRQQLQDLSGESEYKAIQTYLKQTSYRCPN 354
```

```
QY 355 LRHWKVNREGEDRFOAHKSLGNRRLLWHTGNVAVVAAILTSLGRIMPHSGRGVKGII 414
Db 355 LQHIWKVNQGEEDRFOAHKSLGNRRLLWHTGNVAVVAAILTSLGRIMPHSGRGVKGII 414
QY 415 FASNSKSAQYVTMTGCGHGVGMFLGEVALGKEHHITIDDPSSLKSPPPGFSVARGO 474
Db 415 FASNSKSAQYVTMTGCGHGVGMFLGEVALGKEHHITIDDPSSLKSPPPGFSVARGO 474
QY 475 TEPDPAQDIELELDGQPVVPPQPCPSFKSSFSQSEYLYKESQCRRLYLLEIHL 533
Db 475 TEPDPAQDIELELDGQPVVPPQPCPSFKSSFSQSEYLYKESQCRRLYLLEIHL 533

RESULT 3
Q9TX06 ID Q9TX06 PRELIMINARY; PRT; 612 AA.
AC Q9TX06:
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=96007847; PubMed=7578427;
RX Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
SQ SEQUENCE 612 AA; 69241 MW; CB340F7A88FF2364 CRC64;

Query Match 29.2%; Score 822.5; DB 5; Length 612;
Best Local Similarity 38.3%; Pred. No. 3.6e-58;
Matches 209; Conservative 79; Mismatches 207; Indels 51; Gaps 17;

QY 6 KASVQTEGSKKQKQTEEDSFRSTAEALRAAPADNRVIRVDPSCPSFRNPGIOVHED-- 63
Db 96 EATKTAASDLDDSSSESEDEKNOISVKIKGRAAN-----DPHFPSR---XHYENGK 146

QY 64 --YDCTLNQTNIGNNNKFIYIQLLE--EGSRFCWNRWGRVGEVQSGKMHF--TCLEL 117
Db 147 DVIDATLNQTEIQQNNKXIIQLLEADGGSSYVWNRWREGLGKSSSRKDFGKGLNQ 206

QY 118 AKKDFKKFKWEKTKNWEERDRFVAQPNKYTLLEVOGEAESQAVKALSPQVDSGPVRT 177
Db 207 AISLFCSKFEKTKNTFTDRANFKVAGKYDMLDYSTDSPKPK--NGASTATTATTTTK 264

QY 178 VV---RPSCLDPATQNLITNIFSKEMFNAMTLMNLDVKKMPLGKLTQKIARGFEALEA 234
Db 265 VBEHKKECSLDERVQELVLFVFKMMERTMTAKYDLKKMPLGKSKNOITGKYLVLQ 324

QY 235 LEAMKNPTGDGQSLSELSCTFTVTPHNFGRSRPPINSPDVLQAKKMDLLADIELA 294
Db 325 IEDVMGKS--GESLSTLSSRFYTIIPAHFGMSVPPVINTNQMLIEKMNMQLADIEIA 382

QY 295 QTLQAAPEEERKEVEEVPPLDRDYQLLRQQLQDLSGESEYKAIQTYLKQTSYRCPN 354
Db 383 TNLTKDSEDSNI-----LELHYAKLKTDIQPLDENSCYKNNILYVNTYQGGKKPT 436

QY 355 LRHWKVNREGEDRFOAHKSLGNRRLLWHTGNVAVVAAILTSLGRIMPH---SGRVG 410
Db 437 IVNIFKIDRDEADRYKTKKHLGNRRLLWHSRLTNVYASIIISQGLRIAPPEAPVSGYRG 496
```

```

Qy 411 KGIYFASNSKAGVYTTMHCGHGVY-MFLGEVALGKEHHITID---DPSLKSPPPGF 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 497 KGVYFADCMSSLNSCRVTV---GFLIDFCMLLGDVALGK-----TADLARDIYMEKPPQNS 549

Qy 467 DSVIARTGOTEPDPAQDIELELDGQPVVVPQPPVQCPSP-FKSSFSOSEYLIYKESQRL 525
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 HSTWALGVPEDPKPVFFO---DTEGCTIPYQGM---PSQHRGVSCYEHQYVVYDVAQVHL 604

Qy 536 RYLLEI 531
    |||||
Db 605 KYLLQL 610

RESULT 4
O50017 PRELIMINARY; PRT; 653 AA.
AC O50017;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE POLY(ADP-RIBOSE) POLYMERASE.
GN PARP.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OC NCBI_TaxID=4577;
RN [1]
RN SEQUENCE FROM N.A.
RA Babyichuk E., Cottrell P., Storozhenko S., Fuangthong M.,
RA O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT "Higher plants possess two poly(ADP-ribose) polymerases.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ222588; CAAL0888.1;
DR HSSP; P26446; IA26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR003034; SAP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF02037; SAP; 2.
DR SMART; SM00513; SAP; 2.
SQ SEQUENCE 653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

Query Match 25.58; Score 717; DB 10; Length 653;
Best Local Similarity 34.98; Pred. No. 1.5e-49;
Matches 190; Conservative 91; Mismatches 205; Indels 58; Gaps 17;

Qy 13 GSKKOROCETEEDSFRS---TAEALRAAPADNRV---IRVDPSPCFSRNPGIOVHD-Y 64
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 138 GAKEVIKGDDEVEVKKEKWTATKKGAAVLDDQHIPIHRIKYNHYV-----LQVGDEIY 190

Qy 65 DCTLNQTNIQNNNNKFIYIQLLEE--GSRFFCWNRWGRVGVGOSKMHFTCLD-AKKD 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 191 DATLNTQNVGNNNNKFIYIQLVEESDAGSGPMVNRWGRVGVGQDKLHGPSPTRDQAIYE 250

Qy 122 FKFKFWETKKNWEERDRFVAQPNKYTLIEVO-GEAESQAEVVALSPQVDSGPVRTVK 180
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 251 FEGFPHNKTNNHSDRKNFKCYAKKYTWLEMDYGETEK-----EIEKGSITDQIK 300

Qy 181 PCSLDPATONLTNIFSKEMFNAMTLMNLDVKKMPLGKLTQQTARGFEALEEAMK 240
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 301 ETKLETRIAQFISLCINISMKKQWVEIGYNAEKLPGLKRLKATILKGHYHLKRISDVIS 360

Qy 241 NPTGDQGSLEELSSCFYTVIPHNFG--RSRPPPIINSPDVLAQAKDMLLVLAQTLQ 298
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 361 K--ADDRHLEQLTGEFYTVIPHDFGRKMREFIIDTPQKLAKLEWVEALGEIATKLL 418

Qy 299 AAPGEEKVEEVPHPDLDRDQYLLRCQLQLDSESEYKATQYTKQF---GNSYRCNL 355
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 419 -----EDDSSDQDDPLYARYKQLHCOFTPLEADSDSEYSMIKSLRNTHGKSYTVDI 472

```

```

Qy 356 RHVKNVNRGEGRDFOAHKSLGNRRLLWHRGTNAVVAAILTSGLRIMPH-----SGGRVYK 411
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 473 VOIFKVSHEGTERFQKFASTRNRMLLWHGSRLSNWAIGLSQGLRIAPPEAPVTGYMGK 532

Qy 412 GIYFASNSKAGVYTTMHCGHGVY---GYMFLGEVALGKEHHITIDDPSSLKSPPPGDFS 468
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 533 GVIYFADMFESKSANY-----CYASEACRSGLVLLCEVALGDMNELLNADYDANNLPKGLR 587

Qy 469 VIARTGOTEPDPAQDIELELDGQPVVVPQPPVQCPSPFKSSFSOSEYLIYKESQRLRYL 528
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 588 SKGVGTAPNM---VESKVADDGVVVPVLPGEPRQEPS--RGGLLYNEYIYVYNVDQIRMYV 643

Qy 529 LEIH 532
    |||||
Db 644 LHVN 647

RESULT 5
O921K2 PRELIMINARY; PRT; 1014 AA.
AC O921K2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE POLYMERASE).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RN SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012041; AAH12041.1;
KW Transferase.
SQ SEQUENCE 1014 AA; 112721 MW; 5DCE68E4CB3F46EB CRC64;

Query Match 24.98; Score 700.5; DB 11; Length 1014;
Best Local Similarity 33.38; Pred. No. 6.2e-48;
Matches 182; Conservative 100; Mismatches 215; Indels 49; Gaps 17;

Qy 1 MAPKRKASVQTGSGKQKRGTEEDSPRSTAEALRAAPADNRVIRVDPSPCFSRNPGIOV 60
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 494 VAPKGSAAPSKSKS----GAVKEGVNKSERKMKLTUKGGAAYDPDGLSEHSAHLEKG 549

Qy 61 HEDYDCTLNQTNIQNNNNKFIYIQLLEE--SRFFCWNRWGRVGVG--GOSKMHFTCLD 117
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 550 GKVSATLGLVDIVKGTNSYKQLLEDDEKESRYWIFRSWGRVGTIVGSKLEQMPKSED 609

Qy 118 AKKDFKKFWETKKNWEERDRFVAQPNKYTLIEVOGEAESQAEVVALSPQVDSGPVRT 177
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 610 AVEHFMKLYEKTGNAMHSKN-FTKYPKKFYPLEID-YCODEEAVKK-----L 655

Qy 178 VVKP---CSLDPATONLTNIFSKEMFNAMTLMNLDVKKMPLGKLTQQTARGFEALEA 234
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 656 TVKPGTKSLPKPVOELGMIFDVESMKKALVYEIDQKMPGLKSLRQIAAATISLSE 715

Qy 235 LEEAMKNPTGDQGSLEELSSCFYTVIPHNFGSRPPPIINSPDVLAQAKDMLLVLAQTLQ 294
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 716 VQAVSQSSSESQL-DLSNRYFTLIIPHDFGKMPKPLNNADSVQAKVEMDLNLDIEVA 774

Qy 295 QTLQAAPGEEKVEEVPHPDLDRDQYLLRCQLQLDSESEYKATQYTKQF---TGNSY 350
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 775 YSLRGGSDSSK-----DPIDVNYEKLKTDIKVVDROSEAEVIRKYVKNTHATTHAY 829

Qy 351 RCPNLRHVKNVNRGEGRDFOAHKSLGNRRLLWHRGTNAVVAAILTSGLRIMPH-----SG 406
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 830 DL-EVIDIFKIEREGESQRYKPFQLHNRLLWHGSRRTNFAGILUSGLRTAPPEAVTG 888
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Qy 407 GRVKGIFYFASNSKAGVYTTMHCGHGVYMFGEVALGKEHHITIDDPSSLKSPPPGF 466
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```
Db 889 YMFCKGIYFADVMKSYNCHTSQ--GDPIGLILLGEVALGNMYELK-HASHISKLPKKG 945
QY 467 DSVIARGOTEPDPAQDIELELDGQVVPVQGPVQCPSP-FKSSFSQSEVLIYKESQCL 525
Db 946 HSKVGLGKTPDPSASITL-----GVEVPLGTGI--PSGVNDTCLLYNEYIVYDIAQVNL 999
QY 526 YLLEI 531
Db 1000 KYLLKL 1005

RESULT. 6
Q9PS82 PRELIMINARY; PRT; 607 AA.
AC Q9PS82;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSP: P26446; 1A26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR PROSITE: PS50172; BRCT; 1.
SQ SEQUENCE 607 AA; 68033 MW; 75F6EE1D30D8F402 CRC64;

Query Match 24.4%; Score 686.5; DB 13; Length 607;
Best Local Similarity 33.8%; Pred. No. 3.9e-47;
Matches 184; Conservative 97; Mismatches 207; Indels 57; Gaps 20;

QY 5 RKASVQEGSKKQKQ--TEEDSFRSTAEALRAAPADNRVIRVDPCSPFRNPGIQVHED 63
Db 93 KPANMKSAGKVKBEQGPSKSEKKMKLVKGAAVDPDGL--EDSAHVFEKGGKI----- 145
QY 64 YDCTLNQTNIGNNNKYYIIQLLEEG--SRFFCNRWRGVGEV-GQSKMNHFTCLEDAKK 120
Db 146 FSATLGLVDIVKGTNSYKQLLEDDESRVYFRSNGRVGTIGSNKLEQMPKSKEDAVE 205
QY 121 DFKKKFEKTKNKEERDRFVAQPNKTYLIEVOGEAEBSQAVVKALSPQVDSPVTRVVK 180
Db 206 HFLNLYEKTGNSHNSKN-FTKYPKKEYPLEID-YGQDEEAVRKL---TVSAGTKSKLAK 260
QY 181 PCSLDPAQNLITNIFSKEMFKNMTLMNDVKMKPLGKLTQKQIARGFPALEALEAMK 240
Db 261 P-----IQDLIKMIFDVESMKKAMVEFIDLQKMPGLKSKRQIQSAYSILNEVQAVS 314
QY 241 NPTGDDGSLSESCFTVTPHNFGRSRPPPIPSDVLQAKKMDLLVLADIELAQTLOAA 300
Db 315 DGGSESQIL-DLSNRFYTLPHDFGMMKPPLLSNLEYIQAKVQMLDNLDDIEVAYSLLRG 373
QY 301 PGEEKEVEVPPLDRDYOLLCOLQLDSEGESEYKAIQYTKQ-----TGNSYRCPNLR 356
Db 374 GNEDGK-----DPIDINYEKLTDTDKVDKDEEAKIIKQYVNTAATHNAYDL-KVV 427
QY 357 HWKVNREGEGRFQAHSKLGKLNRRLLWHGTNVAVVAAILTSGLRIMPH-----SGGRVKG 412
```

```
Db 428 EIFRIEREGESQRYKPFKQLHNRQLLWHGSRRTTNFAGILLSQGLRIAPPEAPVTGYMFGK 487
QY 413 IYFASENSISAGYVTHMCGGHQGVYMFGEVALGKEH-----HITIDDPKLSKPPPGF 466
Db 488 IYFRDMVMSKSYNCHTSQ--ADPILGILLGEVALGNMYELKNASHIT-----KLPKKG 538
QY 467 DSVIARGOTEPDPAQDIELELDGQVVPVQGPVQCPSPFKSSFSQSEVLIYKESQCLRL 526
Db 539 HSKVGLGKTPADPT--ATTTLDG--VEVPLGNGIS-TGINDTCLLYNEYIVYDVAQVNLK 593
QY 527 YLLEI 531
Db 594 YLLKL 598

RESULT 7
Q9PS81 PRELIMINARY; PRT; 607 AA.
AC Q9PS81;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS Xenopus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae.
OX NCBI_TaxID=8353;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSP: P26446; 1A26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR PROSITE: PS50172; BRCT; 1.
SQ SEQUENCE 607 AA; 67496 MW; 54CDEBBE22079886 CRC64;
```

```
Query Match 24.4%; Score 686.5; DB 13; Length 607;
Best Local Similarity 34.1%; Pred. No. 3.9e-47;
Matches 182; Conservative 95; Mismatches 212; Indels 45; Gaps 19;

QY 10 QTGSKKQKQGT--EEDSFRSTAEALRAAPADNRVIRVDPCSPFRNPGIQVHEDYCTL 68
Db 97 KSSGKVEEKSEKSKMKLVKGAAIDPDS---ELEDSCHVLETGG----KIFSATL 149
QY 69 NOTNIGNNNKYYIIOLLE--EGSRFFCNRWRGVGEV-GQSKMNHFTCLEDAKKDFKKK 125
Db 150 GLVDITFGTNSYKQLIEHDSRIRYVFRSNGRVGTIGSKKLEEMSSDEDAIEHFLNL 209
QY 126 FWEKTKNKEERDRFVAQPNKTYLIEVOGEAEBSQAVVKALSPQVDSPVTRVVKPCSLD 185
Db 210 YQDKTGNWHS-PNFTKYPKKEYPLEI--DYGOEEDVVKLS--VGAGTKSKLAKP---- 260
QY 186 PATQNLITNIFSKEMFKNMTLMNDVKMKPLGKLTQKQIARGFPALEALEBEAKNPTGD 245
Db 261 --VQELIKLIFDVESMKKAMVEFIDLQKMPGLKSKRQIQSAYSILSQVQAVSESLSE 318
QY 246 GQSLSELSSCFTVTPHNFGRSRPPPIPSDVLQAKKMDLLVLADIELAQTLOAAPGEE 305
Db 319 ARLL-DLSNOFYTLPHDFGMMKPPLLNNLEYIQAKVQMLDNLDDIEVAYSLLRGADGG 377
```



Qy 306 EKVEVPHPLDRDYQLLRQQLDSESEYKAIQTYLK-----QTGNSYRCPNLRHVKV 361  
 Db 378 EK-----DPIDVKEKIKTDIKVAKDSESEIRIICDYVKNTHADTHAYDEVL-EIFKI 431  
 Qy 362 NREGEDRFQAHKSLGNRLRHGHTNVAVVAIITSLGLRIMPH-----SGGRVKGIIYFAS 417  
 Db 432 DREGYORYPFKQLHNRQLLHGHGSRRTTNFAGILSQGLRIAPPEAPVGYMGFGKIYFAD 491  
 Qy 418 ENSKSAGVVTVMCGGHQGVGMFGEVALGKEHHTITDDPSLKSPPPGDFSVIARGOTEP 477  
 Db 492 MYSKANYCHAM--PGSPIGLILGEVALGNHKLKASQITKL-PGKHSVKGIGRTAP 548  
 Qy 478 DPAQDIEILDGQVVPVQGPVQCPSPFKSSFSQSEYLIYKESQCRRLYLEI 531  
 Db 549 DPSATV--QLDG--VDVPLGKGTSA-NISDTSLLYNEVIVYDIAQVNLKYLKL 597

## RESULT 8

O81294 ID O81294 PRELIMINARY; PRT; 635 AA.  
 AC O81294;  
 DT 01-NOV-1998 (TrEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
 DE T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE)  
 GN T14P8.19 OR AT4G02390.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA WASHU;  
 RT "The A. thaliana Genome Sequencing Project."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Kalicki J., Elliott G., Cloud J.;  
 RT "The sequence of A. thaliana T14P8."  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Waterston R.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF069298; AAC19283.1; -;  
 DR EMBL; AL161494; CAB80732.1; -;  
 DR HSSP; P26446; 1A26.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR003034; SAP.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF02037; SAP; 2.  
 DR SMART; SM00513; SAP; 2.  
 KW Transferase.  
 SQ SEQUENCE 635 AA; 72017 MW; E3FICBE4D367A377 CRC64;

Query Match 24.3%; Score 683.5; DB 10; Length 635;  
 Best Local Similarity 34.5%; Pred. No. 7.3e-47;

Matches 190; Conservative 90; Mismatches 220; Indels 51; Gaps 19;  
 Qy 2 APKFKASVOTSGSKKQGTDEEDSFRS--TAEALRAAPADNRVIRVDPSPCFPSRNPQIG 59  
 Db 109 APVSSNDKAE--DDNNGFEEKKEKIVTATKGAVALDQWI-----PDEIKSOYHVLO 161  
 Qy 60 VHEIDYDCTLNTQIGNNNKFIYIQLLEGR--FFCWNRWGRVGEVQSKMN-HFTCL 115  
 Db 162 RGDDVDAILNOTNVRNNNKFVQLVLESDSKTKTYMYTRWGRVGVKQSKLDGPDYSW 221  
 Qy 116 EDARKDKFKKFWETKKNKEERDRFVAQPNKYITLIEVOGEAESEAVKALSPQVDSGPV 175  
 Db 222 DRAIEITFNKNDTKNYWSDRKEFIPHKSYTWLEMDYKKEENDSPVNNIPSSS--- 278  
 Qy 176 RTVVKP--CSLDPATQNLITNIFSKEMFKNAMTLNLDVKKMPLGKLPKQOIARGFEALE 233  
 Db 279 --EVKPEQSKLDTRVAKFISLICNVSMMAQHMMEIGYNANKLPGLKISKSTISKGYEVLK 336  
 Qy 234 ALEAMKNPTGDGQSLEELSSCFYTVIPHNFGSRPPD--INSPDVLOAKKMDLLVLADI 291  
 Db 337 RISEVIDR--YDTRLEELSGEFTYVIPHDFGFKKMSQFVIDTPQKLKQKIEMVEALGEI 394  
 Qy 292 ELA--QTLOAAGEEBEKEVEVPHPLDRDYQLLRQQLDSESEYKAIQTYLKQT---G 347  
 Db 395 ELATKLLSVDPGLQDD-----PLYHYQQLNCGLTPVGNDSSEFSMVANYMENTHAKT 447  
 Qy 348 NSYRCPNLRHVKVNRBEGDRFQAHKSLGNRLRHGHTNVAVVAIITSLGLRIMPH--- 404  
 Db 448 HSGVTVEIAQFRASRAVEADRFOOFSKKNRMLLHGHGSRRTTNWAGILSQGLRIAPPEAP 507  
 Qy 405 -SGGRVKGKGIYFASENSKSAGVVTVMHC---GGHQVGMFGEVALGKEHHTITDDPSLK 460  
 Db 508 VTGYMFGKGVYFADMFSSANY-----CYANTGANDGVLLCEVALGDMNELLSDYNAD 562  
 Qy 461 SPPEGFSVIARGOTEPDPAQDIEILDGQVVPVQGPVQCPSPFKSSFSQSEYLIYKE 520  
 Db 563 NLPPGKLSKGVGTAPNPSEAQTL-DG--VVVPLGKPV-E-RSCSKGMLLYNEVIVYNV 618  
 Qy 521 SQCRRLYLEI 531  
 Db 619 EQIKMRYVIQV 629

## RESULT 9

O92P54 ID O92P54 PRELIMINARY; PRT; 983 AA.  
 AC O92P54;  
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).  
 GN PARP-1.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. LANDSBERG RECTA;  
 RA Doucet-chabeau G., Kazmaier M.;  
 RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ131705; CAAL0482.1; -;  
 DR HSSP; P26446; 1A26.  
 DR Transferase; Glycosyltransferase; NAD.  
 FT CHAIN 2 983 POLY(ADP-RIBOSE) POLYMERASE.  
 SQ SEQUENCE 983 AA; 111232 MW; 468E12A8EF1B6F4F CRC64;

Query Match 22.1%; Score 621.5; DB 10; Length 983;  
 Best Local Similarity 32.7%; Pred. No. 1.5e-41;  
 Matches 179; Conservative 87; Mismatches 209; Indels 73; Gaps 21;





```
AC 09ZSV1.
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN PARP1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99026291; PubMed=9808734;
RA Mahajan P.B., Zuo Z.;
RT "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL Plant Physiol. 118:895-905(1998).
DR EMBL; AF093627; AAC79704.1; -.
DR HSSP; P26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR003034; SAP_reg.
DR InterPro; IPR001510; ZnF-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; ZnF-PARP; 2.
DR ProDom; PD004675; ZnF-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS50172; BRCT; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;
```

```
Query Match      20.8%; Score 584.5; DB 10; Length 980;
Best Local Similarity 30.7%; Pred. No. 1.6e-38;
Matches 164; Conservative 98; Mismatches 199; Indels 73; Gaps 20;

QY 33 ALRAAPADNRVIRVDSPCFNRNCIQ-----VHED----YDCTLNQTNIGNNKFFYIIQ 84
DB 479 ALESKSGTGVTVKGRSAVHSSGLOQTAHLEDKSIYNATLNMSDLALGVNSYVYLQ 538
QY 85 LLE--EGSRFFCWNKWRGV--EYQSKMNHFTCLEDAKKDFKKFKWEKTKNKWEE---R 137
DB 539 IIEQDDGSECYVFRKWRGVSGEKIGGKLEMSKTE-AIKEFKRLFLEKTSWEAWECK 597
QY 138 DRFVAQPNKYTLIEVQGEASEAVVKALSPQVDSGPVTVVVKPCSLDPATQNLITNIFS 197
DB 598 TNFRKQPGRFPLDV-----DYGKKAPKRKDISEMKS-----SLAPQLLELMKMLFN 645
QY 198 KEMFNAMTLNLDVKKMPLGKLTQQTARCFEALAELEAMKNPTGDGQSLER-----L 252
DB 646 VETRYAAMHEFEINNSEMPGLKSKENIEKGFALTEIQNLKDTADQALAVRESLIVAA 705
QY 253 SSCFTYVIPHNGRSRPPINSPDVLQAKDKMLLVADIQAOTLQAAPEEGEEKVEVP 312
DB 706 SNRFTFLIP-----SHPHIIRDEDDLMKAKMLEALQDIETASKIVGFDSDSES----- 756
QY 313 HPLDRDYQLLCQLQLDGSSEYKATQTYLKTGNSYRCP-----NLRHWKVNREG 365
DB 757 --LDDKVMKLCDDITPLAHDSEDKLIEQYLLNT-----HAPTHKDSWSELEEVSLDRG 810
QY 366 EGDGFQAH-SKLGNRRLLWHGNTNVAVAAIILTSGLRIMPH-----SGRVGKGIYFASNS 420
DB 811 ELNYSRYKNLHKNMLLHWSGRITNFVGLISQGLRIAPPAPVGYMGKGLYFADLVS 870
QY 421 KSAQYVTTMHC---GGHQVGYMFLGEVALGKEHHITIDDPGLKSPPPGFSVITARGOTEP 477
DB 871 KSAQY-----CYVDRNPVGLMLLSEVALGDMYELK-KATSMDKPPRGKHSKGLGKTPV 924
QY 478 DPAQDIELEDGQPVVVPVQGPVQCPQCFKSSFSQSEYLYIKESQCRLLYLEI 531
```

```
DB 925 LESEFVKWRDD---VVVPCGKVPV--SSIRSELMYNEYIVNTSQVKMQFLKV 974

RESULT 14
Q9N4H4
ID Q9N4H4 PRELIMINARY; PRT; 945 AA.
AC Q9N4H4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 108.0 KDA PROTEIN.
GN Y71F9AL.18.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RX None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Bradshaw-Cordum H., Scott K., Graves T.;
RT "The sequence of C. elegans cosmid Y71F9AL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC024200; AAC36011.1; -.
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR001510; ZnF-PARP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF00645; ZnF-PARP; 1.
DR ProDom; PD004675; ZnF-PARP; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
KW Hypothetical protein
SQ SEQUENCE 945 AA; 108006 MW; 1D0A62C954BC6AD9 CRC64;
```

```
Query Match      20.2%; Score 569.5; DB 5; Length 945;
Best Local Similarity 31.5%; Pred. No. 2.4e-37;
Matches 156; Conservative 80; Mismatches 201; Indels 59; Gaps 14;

QY 64 YDCTLNQTNIGNNKFFYIIQLLEGSR--FFCWNKWRGV--EYQSKMNHFTCLEDAKK 120
DB 477 YQATLSFTDLTONKSYKIQLLDKDDQRRYVFRSGRVGTEVGNKHSYSNSNAIL 536
QY 121 DFKKKFKFKTKNWEERDRFVAQPNKYTLIEVQGEASEAVVKALSPQVDSGPVTVVK 180
DB 537 KFQDVFEKTKNDWIYRKHKFKMFGMSYVETDYSEFAQ-----ITDTEIT 582
QY 181 PCS---LDPATQNLITNIFSKEMFNAMTLNLDVKKMPLGKLTQQTARCFEALAELE 237
DB 583 PGSKTLPLKSVKVEVMVIFDVENKMSALKSFEMDVNKMPLGRLSHNOINLAFAVLNDISD 642
QY 238 AMKNPTGDGQSLSELSCTYTVIPHNFGSRPPINSPDVLQAKDKMLLVADIQAOTL 297
DB 643 LLVKLPLDASRLIDFSNKFTYIIPHNFGMRVPEIDSFHKIKKNNMLNALLDIKFAYD- 701
QY 298 QAAPGEEEEKVEVPHPDLDRDYQLLCQLQLDGSSEYKATQTYLKTGNSYRCP-NL 355
DB 702 QISGGDVPASTSLGIDPVDINYOGLKCIEMPELQGGCDDWNHIOYLANTHGATHDLKVEL 761
```

```
Qy 356 RHVWVNRREGDRFOAHSKLGRRLLWHGNTNVAVAAILTSGLRIMPH-----SRRVGVK 411
Db 762 IDILKLRDNSSKFKRH--IGNRLLWHGSKRMNFAGILGQGLRIAPPEAPVSGYMGK 819
Qy 412 GIYFASNSKSGAGVYVTTMHCGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVI 470
Db 820 GYFADMFSSFFY---CRANAKEEAYLLLCDAVALGNVQQLMASKNVSROTLPAGFQSVQ 876
Qy 471 ARGQ-----TEPD-----PAQDIELELDGQPVVVPOGPPVQCPGSKSSFSQSEYLI 517
Db 877 GLGRQCPREIGSYNHPDGYTIPGLTYMQLQKQDV-----DYHLLYNEFIV 923
Qy 518 YKESQCRRLRYLLEIHL 533
Db 924 YVDQIQIKLYLRVKM 939

RESULT 15
Q9XUA5 PRELIMINARY; PRT; 727 AA.
AC Q9XUA5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AC8.1 PROTEIN.
OS AC8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83097; CAB05448.1; -.
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; zf-PARP; 1.
DR ProDom; PD004675; Znf-PARP; 1.
DR ProSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFEDF CRC64;
```

Query Match 16.9%; Score 475.5; DB 5; Length 727;  
Best Local Similarity 30.2%; Pred. No. 7e-30;  
Matches 134; Conservative 68; Mismatches 170; Indels 71; Gaps 13;

```
Qy 126 FWEKTKNKEERDRFVAQPNKYTLIEVQGE--AESQEAIVKALSPQVDSGPRVTVKPCS 183
Db 315 FHEKTKNDIYRKHFRKMPGMFYSVETDYSEFVGTNNGHKKKITPGSKITPGSKTLLPKS 374
Qy 184 LDPATONLTINIFSKEMFKNMTLMLNDVKKMPILKTKQIARGFEALEALEAMKNPT 243
Db 375 V-----KEVMSIFDVENKSKALKSFEIDVKNMPLGRUSHNQINLAFEVLENDISDLLVCLKP 430
Qy 244 GDGQSLSELSCEFTYVIPHNGRSRPPPIINSPDVLOAKDKMLLVADIETLAQTLQAAPGE 303
Db 431 IDASKILDFSNKFTYTIIPHNGRVPPEPIDSFHKIKEKNMNLALLDIKFAYD-QISGGD 489
Qy 304 EEKVEEVPHPLDQRYOLLRCQLQLDGSESEYKAQTLYLKOT-GNSYRCP-NLRHWVKY 361
Db 490 VPASTSLSIDPVDINRYKLKCIIMEPLQGGDDWNMIHOYLKNTHGATHDLKVELIDILKV 549
```

```
Qy 362 NREGGDRFOAHSKLGRRLLWHGNTNVAVAAILTSGLRIMPH-----SRRVGVKGIYFAS 417
Db 550 NRDNSSKFKRH--IGNRLLWHGSKRMNFAGILGQGLRIAPPEAPVSGYMGKGYFAD 607
Qy 418 ENSKSAGYVYVTTMHCGHQVGYMFLGEVALGKEHHITID-DPSLKSPPPGFDSVIARGOTE 476
Db 608 MFSKSFY---CRANAKEEAYLLLCDAVALGNVQQLMASKNVSROTLPAGFQSV----- 657
Qy 477 PDPAQDIELELDGQPVVVPOGPPVQCP-----SFKSS-----SFSQ----- 512
Db 658 -----QGVGRQCPREIGSYNHPDGYTIPGLTYMQLQKQKQNVYHL 698
Qy 513 --SEYLIYKESQCRRLRYLLEIHL 533
Db 699 LYNEFIYVDVQIQIKLYLRVKM 721
```

Search completed: August 29, 2002, 08:01:37  
Job time: 369 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run On: August 29, 2002, 07:59:24 ; Search time 69.02 Seconds  
(without alignments)  
742.040 Million cell updates/sec

Title: US-09-701-586b-4

Perfect score: 2823

Sequence: 1 MAPKPKPWQTEGPEKKGR.....EYLIYQESQCRRLYLEVHL 533

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71.\*

1: pir1.\*

2: pir2.\*

3: pir3.\*

4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2419	85.7	459	2 T08713	NAD+ ADP-ribosyltr
2	694	24.6	635	2 T01311	NAD+ ADP-ribosyltr
3	677	24.0	1016	1 J50428	NAD+ ADP-ribosyltr
4	674.5	23.9	996	1 S42208	NAD+ ADP-ribosyltr
5	670.5	23.8	653	2 T03656	probable NAD+ ADP-
6	669.5	23.7	1011	1 JH0581	NAD+ ADP-ribosyltr
7	666.5	23.6	1014	1 A29725	NAD+ ADP-ribosyltr
8	657	23.3	500	2 S26057	NAD+ ADP-ribosyltr
9	653	23.1	998	2 S17335	NAD+ ADP-ribosyltr
10	649.5	23.0	1013	1 S04200	NAD+ ADP-ribosyltr
11	638	22.6	994	1 A47474	NAD+ ADP-ribosyltr
12	598.5	21.2	983	2 T51353	NAD+ ADP-ribosyltr
13	598.5	21.2	1009	2 C84719	probable poly (ADP
14	578	20.5	969	2 T03657	NAD+ ADP-ribosyltr
15	534	18.9	727	2 T18600	hypothetical prote
16	393.5	13.9	538	2 T20414	hypothetical prote
17	190.5	6.7	135	2 P80494	NAD+ ADP-ribosyltr
18	164.5	5.8	2004	2 D88948	protein ZK1005.1 (
19	134	4.7	1156	2 B70356	chromosome assembl
20	126	4.5	1222	2 C88504	protein B0361.3 [i
21	120.5	4.3	2044	2 A81180	probable peptidogl
22	120	4.3	1938	1 A40997	myosin heavy chain
23	116.5	4.1	1435	2 S69632	regulatory protein
24	113.5	4.0	181	2 T3058	NAD+ ADP-ribosyltr
25	113.5	4.0	1134	2 A60234	Iga Fc receptor pr
26	113.5	4.0	1164	1 FCSOAG	Iga Fc receptor pr
27	113.5	4.0	4540	2 T30838	cytoplasmic dynein
28	110.5	3.9	2078	2 T25400	hypothetical prote
29	110	3.9	1534	2 A56734	ribosome receptor,

30 110 3.9 2288 2 T29999  
31 109.5 3.9 446 1 A46335  
32 108.5 3.8 388 1 QOYV  
33 108.5 3.8 1233 2 T30534  
34 108.5 3.8 1871 2 S27938  
35 108 3.8 359 2 T12540  
36 108 3.8 1092 2 T33717  
37 108 3.8 3225 2 I52300  
38 108 3.8 3259 1 A56539  
39 107.5 3.8 716 2 T26998  
40 107 3.8 1186 2 G69708  
41 106.5 3.8 568 2 T34522  
42 106 3.8 725 1 JC5016  
43 106 3.8 932 2 S62555  
44 106 3.8 1199 2 T29145  
45 106 3.8 2176 2 T13806

#### ALIGNMENTS

RESULT 1

T08713

NAD+ ADP-ribosyltransferase homolog DKFzp566G0224.1 - human (fragment)

C:Species: Homo sapiens (man)

C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 03-Nov-2000

C:Accession: T08713

R:Ansorge, W.; Wirkner, U.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

Submitted to the Protein Sequence database, May 1999

A:Reference number: Z16472

A:Accession: T08713

A:Molecule type: mRNA

A:Residues: 1-459 <ANS>

A:Cross-references: EMBL:AL050034

A:Experimental source: fetal kidney; clone DKFzp566G0224

C:Genetics:

A:Note: DKFzp566G0224.1

C:Superfamily: NAD+ ADP-ribosyltransferase

Query Match 85.7%; Score 2419; DB 2; Length 459;  
Best Local Similarity 99.8%; Pred. No. 6.3e-162;  
Matches 458; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 75 IENNKKFYIIQLQDSNRRFFTCNNRWGRVGEVGSQKINHFTREDAKKDFEKKERKTK 134

Db 1 IENNKKFYIIQLQDSNRRFFTCNNRWGRVGEVGSQKINHFTREDAKKDFEKKERKTK 60

Qy 135 NNWAERDHFVSHPGKYTLIEVQAEDEAQAQVAVKVDGPGVVRTVTKRQPCSLDPATQKLT 194

Db 61 NNWAERDHFVSHPGKYTLIEVQAEDEAQAQVAVKVDGPGVVRTVTKRQPCSLDPATQKLT 120

Qy 195 NTFSEMFKNMTALMDLVKKMPLGKLSKQQTARGFEEALEEALKGPTDGGQSLEELS 254

Db 121 NTFSEMFKNMTALMDLVKKMPLGKLSKQQTARGFEEALEEALKGPTDGGQSLEELS 180

Qy 255 SHFYTVIPHNFGHSOPPPINSPELLQAKKDLLVLADIQAQAVSQEKTVEVPHP 314

Db 181 SHFYTVIPHNFGHSOPPPINSPELLQAKKDLLVLADIQAQAVSQEKTVEVPHP 240

Qy 315 LDRDYLKLCQQLQDSGAPEYKVTQTYLEQTSNHRCPQLQHIWKNVGEDEDFQAH 374

Db 241 LDRDYLKLCQQLQDSGAPEYKVTQTYLEQTSNHRCPQLQHIWKNVGEDEDFQAH 300

Qy 375 KLGNRKLWHGNTNMAVVAAILTSGLRIMPHSGRGVKGKGIYFASENSKSAGYVGMKCGAH 434

Db 301 KLGNRKLWHGNTNMAVVAAILTSGLRIMPHSGRGVKGKGIYFASENSKSAGYVGMKCGAH 360

Qy 435 HVGVMFLGEVALGREHHINTDNPSLKSPPPGDSVIARTGHTPDPTQDTELELDGQVVV 494

Db 361 HVGVMFLGEVALGREHHINTDNPSLKSPPPGDSVIARTGHTPDPTQDTELELDGQVVV 420

Qy 495 PQGQVPVCPPEFSSSITSQSEYLIYQESQCRRLYLEVHL 533

hypothetical prote  
gag polyprotein -  
transforming prote  
chromosome segrega  
hypothetical prote  
hypothetical prote  
carbamoyl-phosphat  
giantin - human  
giantin - human  
hypothetical prote  
chromosome segrega  
hypothetical prote  
hyaluronan recepto  
protoplast regener  
hypothetical prote  
toucan gene protei

```

QY 473 GHTPDPDTQTELDCCQVVVFGQPV--PCPEFSSSTFSQSEYLIYQBSQRLRYLLE 530
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 575 GKTAPNPSEAQTLE-DG--VVVPLGKPVRSCL---SKGMLLYNEYIVYNVEQIKMRYVIQ 628

QY 531 V 531
Db 629 V 629

RESULT 3
JS0428
NAD+ ADP-riboseyltransferase (EC 2.4.2.30) - bovine
N/Alternate names: ADP-riboseyltransferase (polymerizing); poly(adenosine diphosphate
C/Species: Bos primigenius taurus (cattle)
C/Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 19-Jan-2001
C/Accession: JS0428; S00328; A30458
R:Saito, I.
submitted to JIPID, February 1990
A/Reference number: JS0428
A/Accession: JS0428
A/Molecule type: mRNA
A/Residues: 1-1016 <SAI>
A:Experimental source: thymus
R:Taniguchi, T.; Yamauchi, K.; Yamamoto, T.; Toyoshima, K.; Harada, H.; Tanaka, H.; T
Eur. J. Biochem. 171, 571-575, 1988
A/Title: Depression in gene expression for poly(ADP-ribose) synthetase during the int
A/Reference number: S00328; MUID:88151954
A/Accession: S00328
A/Molecule type: mRNA
A/Residues: 648-714; 838-904 <TAN>
A/Cross-references: EMBL:X06986
A/Accession: A30458
A/Molecule type: protein
A/Residues: 658-685; 689-696; 893-901 <TA2>
C/Superfamily: NAD+ ADP-riboseyltransferase

```

F; 21-51/Region: zinc finger  
F; 128-165/Region: zinc finger  
F; 200-230/Region: helix-turn-helix motif  
F; 224-231/Region: helix-turn-helix motif  
F; 250-270/Region: helix-turn-helix motif  
F; 494-501/Region: nucleotide-binding motif A (P-loop)  
F; 890-903/Region: nucleotide binding #status predicted

Query Match 24.0%; Score 677; DB 1; Length 1016;  
Best Local Similarity 33.8%; Pred. No. 2.8e-39;  
Matches 187; Conservative 99; Mismatches 205; Indels 62; Gaps 24;

QY 1 MAYPKP---PWVQTEGPEKKKGROAGREDPPRSTAEALKAIPAERRIIRVDPTCLSS 56  
Db 495 VGPKGSGAAPSKKSGKPVKEEG--TNKSKRMKLTKGAA-----VDPDSGLEH 543  
QY 57 NPGTVQVED----YNCNLNTNIENNNNKRYIYLQLQDSNRF-FTCWNRMGRVGEY-GOS 110  
Db 544 N--AHVLEKGGKVFESATLGLVDIVKGTNSYYKLQLEDDKESRYWIFRSGRVGTVIGSN 601  
QY 111 KINHFTLEDAKKDFEKKFKREKTNWAERDHFVSPHGKTYLIEVQ-AEDEAQEAENVKVD 169  
Db 602 KLEQMPSSKDAIEHFMKLYEKTGTNAWHSK-NFTKHPKFPYLEIDYGOE--EAVKKLT 658  
QY 170 RGPVRTVTKRVQCSLDPA\*QKLITNIFSKEMFNKTMALMDLVKMKPLGKLSKQOIARG 229  
Db 659 VNP-GTKSKLPK-----VQNLIKMFDSVMKKAWEYEDLQKMPLGKLSKROIQA 711  
QY 230 FEALEALEAL-KGPTDGGOSLEELSSHFTVTYPHNFGHSQPPINSPELLOAKKDWLLV 288  
Db 712 YSILSEVQQALSGDSSD---SHIIDLNSRFTYLLPHDFGMKKPPLLNANNSVQAKWELON 769  
QY 289 LADIELAQAALQAVSEQEKTVEEYPHPLDRDYQLLKCOLQLDLSGAPYKVIQTYLEQT-G 347  
Db 770 LLDIEVAYSLRGSDSDSSKD---PIDVNEKLTDTIKVDDKDSSEAEIRKYVKNTHA 825





C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C: Accession: JH0581  
R: Ittel, M.E.; Garnier, J.M.; Jeltsch, J.M.; Niedergang, C.P.  
Gene 102, 157-164, 1991  
A: Title: Chicken poly(ADP-ribose) synthetase: complete deduced amino acid sequence and cDNA  
A: Reference number: JH0581; MUID: 91340148  
A: Accession: JH0581  
A: Molecule type: mRNA  
A: Residues: 1-1011 <ITP>  
A: Cross-references: EMBL:X52690; NID:G63742; PIDN:CAA36917.1; PID:G63743  
C: Comment: This protein is a chromatin-bound enzyme.  
C: Comment: This enzyme catalyzes DNA-dependent post-translational modifications of various proteins.  
C: Superfamily: NAD+ ADP-ribosyltransferase  
C: Keywords: DNA binding; glycosyltransferase; NAD; nucleus; pentosyltransferase; zinc finger

Query Match 23.7%; Score 669.5; DB 1; Length 1011;  
Best Local Similarity 33.2%; Pred. No. 9.5e-39;  
Matches 186; Conservative 100; Mismatches 205; Indels 69; Gaps 24;

QY 7 PW---VQTECP-----KKGRQAGREEDPFRSTAFALKAIPAEXRIIRVDPT 51  
DB 477 PWGAKEVTHEQVADVGVKCKPANMSAGKVEEQGPKSEKMKLTV---KGGAAVDDP 533  
QY 52 CPLSSNPGTQVYED---NYCTLNQNTNIENNKKFYIIQLQD---SNRFFTCWNRGRVG 105  
DB 534 SGLDS--AHVEFGKGFISATLGLVDIVKGTNSYKQLLEDDEDRSYVW-FRSWGRVG 590  
QY 106 EV-GQSKINHFTLEDAKDFEKKFKYKNNAERDFVSHPGKYLIEVQ-AEDEAQE 163  
DB 591 TVTGSNKLEQMPKSDAVEHFLNLYEKTGNSWHSK-NFTKPKFYPLEIDYQDDEAV 649  
QY 164 AVVKVDRGVPVTRTKRVPCLDPATQKLTITNIFSKEMFKNTMALMDLDVKKMPLGKLSK 223  
DB 650 RLTVSAG---TKSLAKP-----IQDLKMFVDSKMKKAMVEFIDLQKPLGKLSK 700  
QY 224 QQTARGEALAEALKAALGPTDGGOSLE--ELSSHFTYVTPHNFGRSPPPIINPELLQA 281  
DB 701 RQIQASVILNEVQAV---SDGSESQILDSNREYTLPHDFGKMKPPLLSNLEYIOA 757  
QY 282 KDMLLVLADIELAQAQVSEQKTVVEVPHPLDRDYOLLKCOLQLDLSGAPEYKVIQT 341  
DB 758 KVQMLDLIDIEVAYSLRGNEDGDKD---PIDINYEKLFTDIKVDKDSSEAKIIO 813  
QY 342 YLEQT-GSNHRCPTLQ--HTWKVNGEEDRFQAHSKLGRKLLHGTNNMVAAILTSG 398  
DB 814 YVNTAATHANVADLKVVEIFRIEREGESQRYKPFQQLHNRQLLHWSRTNFGILSOG 873  
QY 399 LRIMPH-----SGRVGKGYFASNSKSGAGYVIGMKC---GAHVGYMFLGEVALGREHH 451  
DB 874 LRAPPEAPVTGYMFGKGYFRDMVSKSANY-----CHTSQADPIGLILLGEVALGNMVE 928  
QY 452 INTDNLSPKSPFPFDSVATRGHTEDPTQDTELDGQGVVQVQPCVPCPEFSSSTFS 511  
DB 929 LKNASHITKL-PKGKSVAGLGTAPDPTATT--TLDG--VEYPLNGI-STGINDTCLL 982  
QY 512 QSEYLIQESOCRLRYLLEV 531  
DB 983 YNEYIYDVDAQVNLKLLKL 1002

RESULT 7  
A29725  
NAD+ ADP-ribosyltransferase (EC 2.4.2.30), nuclear - human  
N: Alternate names: poly (ADP-ribose) polymerase; poly (ADP-ribose) synthetase; poly(ADP)  
C: Species: Homo sapiens (man)  
C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C: Accession: A29725; A28498; A39976; A26901; I38096; B33321; A35635; A61559; S14  
R: Ochiai, K.; Morita, T.; Sato, T.; Ogura, T.; Yamashita, R.; Noguchi, S.; Suzuki, H.; N  
Biochem. Biophys. Res. Commun. 148, 617-622, 1987  
A: Title: Nucleotide sequence of a full-length cDNA for human fibroblast poly(ADP-ribose)  
A: Reference number: A29725; MUID: 88076933  
A: Accession: A29725

A: Molecule type: mRNA  
A: Residues: 1-69, 'O', '71-1014 <UCH>  
A: Cross-references: GB:M18112; NID:G190166; PIDN:AAA60137.1; PID:G190167  
R: Kurosaki, T.; Ushiro, H.; Mitsuuchi, Y.; Suzuki, S.; Matsuda, M.; Matsuda, Y.; Kato  
J. Biol. Chem. 262, 15990-15997, 1987  
A: Title: Primary structure of human poly (ADP-ribose) synthetase as deduced from cDNA  
A: Reference number: A28498; MUID: 88058958  
A: Accession: A28498  
A: Molecule type: mRNA  
A: Residues: 1-16, 'E', '18-211, 'K', '213-236, 'R', '238-366, 'H', '369-1014 <KUR>  
A: Cross-references: GB:J03473  
R: Cherney, B.W.; McBride, O.W.; Chen, D.; Alkhatib, H.; Bhatia, K.; Hensley, P.; Smul  
Proc. Natl. Acad. Sci. U.S.A. 84, 8370-8374, 1987  
A: Title: cDNA sequence, protein structure, and chromosomal location of the human gene  
A: Reference number: A39976; MUID: 88068596  
A: Accession: A39976  
A: Molecule type: mRNA  
A: Residues: 1-49, 'D', '51-612, 'Q', '614-907, 'Y', '909-939, 'R', '941-979, 'I', '981-1014 <CHE>  
A: Cross-references: GB:J03030  
A: Note: The authors translated the codon ATA for residue 980 as Asn  
R: Suzuki, H.; Uchida, K.; Shima, H.; Sato, T.; Okamoto, T.; Kimura, T.; Miwa, M.  
Biochem. Biophys. Res. Commun. 146, 403-409, 1987  
A: Title: Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and expressi  
A: Reference number: A26901; MUID: 87298455  
A: Accession: A26901  
A: Molecule type: mRNA  
A: Residues: 441-610, 'N', '612-880, '921-1014 <SUZ>  
A: Note: The sequence figure has an omission of forty residues  
R: Ogura, T.; Niyunoya, H.; Takahashi-Masutani, M.; Miwa, M.; Sugimura, T.; Esumi, H.  
Biochem. Biophys. Res. Commun. 167, 701-710, 1990  
A: Title: Characterization of a putative promoter region of the human poly(ADP-ribose)  
A: Reference number: I38096; MUID: 90211250  
A: Accession: I38096  
A: Status: translation not shown  
A: Molecule type: DNA  
A: Residues: 1-40 <RES>  
A: Cross-references: EMBL:X16674; NID:G510112; PIDN:CAA34663.1; PID:G1017423  
R: Auer, B.; Nagl, U.; Herzog, H.; Schneider, R.; Schweiger, M.  
DNA 8, 575-580, 1989  
A: Title: Human nuclear NAD(+) ADP-ribosyltransferase (polymerizing): organization of t  
A: Reference number: A33321; MUID: 90091744  
A: Accession: B33321  
A: Molecule type: DNA  
A: Residues: 38-43, '93-98; 132-137, '204-209; 237-242; 276-281; 335-340; 384-389; 431-436; 512-5  
A: Cross-references: GB:M29544; GB:M22953  
A: Note: The authors translated the codon GTG for residue 54 as Glu  
A: Note: these fragments represent intron-exon boundaries  
A: Accession: A33321  
A: Status: nucleic acid sequence not shown; not compared with conceptual translation  
A: Molecule type: DNA  
A: Residues: 16-66; 96; 121-159, 'D', '161-167 <AU2>  
A: Note: these fragments represent a zinc finger-containing DNA-binding region  
R: Gradwohl, G.; Menissier de Murcia, J.; Molinete, M.; Simonin, F.; Koken, M.; Hoeijm  
Proc. Natl. Acad. Sci. U.S.A. 87, 2990-2994, 1990  
A: Title: The second zinc-finger domain of poly(ADP-ribose) polymerase determines spec  
A: Reference number: A35635; MUID: 90222155  
A: Accession: A35635  
A: Status: not compared with conceptual translation  
A: Molecule type: mRNA  
A: Residues: 12-26, 'T', '28-66; 116-166 <GRA>  
R: Schneider, R.; Auer, B.; Kuehne, C.; Herzog, H.; Klocker, H.; Bartscher, H.J.; Hirs  
Eur. J. Cell Biol. 44, 302-307, 1987  
A: Title: Isolation of a cDNA clone for human NAD (+): protein ADP-ribosyltransferase.  
A: Reference number: A61559; MUID: 88082900  
A: Accession: A61559  
A: Molecule type: mRNA  
A: Residues: 381-420; 682-710 <SCH>  
R: Yokoyama, Y.; Kawamoto, T.; Mitsuuchi, Y.; Kurosaki, T.; Toda, K.; Ushiro, H.; Tera  
Eur. J. Biochem. 194, 521-526, 1990  
A: Title: Human poly(ADP-ribose) polymerase gene. Cloning of the promoter region.  
A: Reference number: S14010; MUID: 91099327  
A: Accession: S14010  
A: Status: preliminary









Db 648 QTLPAQSQVQGVGRQCPREIGSYKNPDGTYVPLGLITYMLOGKQNV----- 694  
QY 507 SSTFSQSEYLIYQESQCRRLRYLLEVHL 533  
Db 695 DYHLLYNEFIVYVDQIQLKYLVRVKM 721

Search completed: August 29, 2002, 07:59:29  
Job time: 276 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:58:18 ; Search time 46.79 Seconds  
(without alignments)  
278.240 Million cell updates/sec

Title: US-09-701-586b-8

Perfect score: 2813

Sequence: 1 MAPKKASVQTEGSKKQKQ.....EYLIYKESQCLRYLLEIHL 533

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCRUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	705	25.1	1013	4	US-08-860-886-2
2	700	24.9	1014	4	US-09-078-347A-3
3	146	5.2	1327	4	US-09-196-387-2
4	105.5	3.8	635	4	US-09-046-992-2
5	104	3.7	2482	1	US-08-328-254-6
6	103.5	3.7	613	1	US-08-405-615-1
7	103.5	3.7	613	2	US-08-461-234-1
8	103.5	3.7	613	2	US-08-463-480-1
9	103.5	3.7	614	3	US-08-225-224-1
10	103.5	3.7	614	3	US-08-722-258-1
11	103.5	3.7	614	5	PCT-US95-04468-1
12	103.5	3.7	638	3	US-09-047-148-2
13	100.5	3.6	420	1	US-08-391-259-7
14	100.5	3.6	420	2	US-08-839-425-7
15	100.5	3.6	889	4	US-09-336-447A-15
16	99.5	3.5	462	2	US-08-484-438-42
17	99	3.5	497	2	US-08-511-485-4
18	99	3.5	497	3	US-09-212-971-4
19	99	3.5	497	4	US-08-800-929A-4
20	99	3.5	497	4	US-09-617-053A-4
21	99	3.5	622	2	US-08-356-786-16
22	99	3.5	3248	1	US-08-353-700-1
23	99	3.5	3248	5	PCT-US95-16216-1
24	98.5	3.5	579	1	US-08-126-564A-31
25	98.5	3.5	579	5	PCT-US94-09143-31
26	98	3.5	3075	2	US-08-460-309-5
27	98	3.5	3075	2	US-08-125-077-5

28 97.5 3.5 508 1 US-08-021-601-10 Sequence 10, Appl  
29 97.5 3.5 508 1 US-08-082-849B-10 Sequence 10, Appl  
30 97.5 3.5 508 5 PCT-US94-01624-10 Sequence 10, Appl  
31 97.5 3.5 2285 4 US-09-308-375-2 Sequence 2, Appl  
32 96 3.4 163 1 US-08-044-618-6 Sequence 6, Appl  
33 96 3.4 637 1 US-08-235-838-14 Sequence 14, Appl  
34 96 3.4 637 2 US-08-465-473B-14 Sequence 14, Appl  
35 96 3.4 1128 4 US-08-923-992A-6 Sequence 6, Appl  
36 96 3.4 1164 4 US-08-923-992A-2 Sequence 2, Appl  
37 95.5 3.4 396 4 US-09-046-992-4 Sequence 4, Appl  
38 95.5 3.4 420 1 US-08-391-259-2 Sequence 2, Appl  
39 95.5 3.4 420 1 US-08-391-259-11 Sequence 11, Appl  
40 95.5 3.4 420 2 US-08-839-425-2 Sequence 2, Appl  
41 95.5 3.4 420 2 US-08-839-425-11 Sequence 11, Appl  
42 95.5 3.4 599 1 US-08-463-163-3 Sequence 3, Appl  
43 95.5 3.4 637 1 US-08-235-838-16 Sequence 16, Appl  
44 95.5 3.4 637 2 US-08-465-473B-16 Sequence 16, Appl  
45 95 3.4 816 2 US-08-533-306A-6 Sequence 6, Appl

#### ALIGNMENTS

RESULT 1  
US-08-860-886-2  
; Sequence 2, Application US/08860886  
; Patent No. 6335009  
; GENERAL INFORMATION:  
; APPLICANT: Burkli, Alexander  
; APPLICANT: zur Hausen, Harald  
; APPLICANT: Jan-Heiner, Kupper  
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE  
; TITLE OF INVENTION: IN GENE THERAPY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds, LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/860,886  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 8484-0028-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-08-860-886-2

Query Match 25.1%; Score 705; DB 4; Length 1013;





Db 329 QVQVIRNALASP-GSGDLGEAT-----REQPE-----QAR--LAULTLA 365  
QY 290 DIELAOTLOAAPGEEKEVEVPHPLDRDYQLRCQLQL-----LDSG-----ESEYKAI 339  
Db 366 AASERFVROGTGDEAGA-----ANADVSLTFCVPAAGECAGPADSGDALLLEANTPTG 419  
QY 340 QTYLKOTGNSYRCNLRHVWVKNVREGEDRFQAHSKLGNRRLL---WHGTNVAVVAAILT 396  
Db 420 AEFGLDGDVSVFSTRGTQNTVER-----LLQAHRLQLEERGYYVGVYGFLEAAQSIIVF 474  
QY 397 SGLRIMPHSGRVCCKGIYFASSENSKSGYVVTMCGGHGVGMPLGEVALGKEHHITIDD 456  
Db 475 GGVARSQDLDAIWRGFIYAGDPALAYGYAQDQEPDAR-----GRIRGALLRVVYVPR 527  
QY 457 PSLSPPPPGSDSVIARGTQPTPDPAODIELELDGQPV 492  
Db 528 SSL-----PGFYRT-SLTLAAPEAAAGEVE-RLIGHPL 557

RESULT 5  
US-08-328-254-6  
; Sequence 6, Application US/08328254  
; Patent No. 5710022  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Xueliang  
; APPLICANT: Lee, Wen-Hwa  
; TITLE OF INVENTION: A No. 5710022el Nuclear Mitotic Phosphoprotein  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Campbell and Flores  
; STREET: 4370 La Jolla Village Drive, Suite 700  
; CITY: San Diego  
; STATE: California  
; COUNTRY: USA  
; ZIP: 92122  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/328,254  
; FILING DATE: 24-OCT-1994  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/141,239  
; FILING DATE: 22-OCT-1993  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Campbell, Cathryn A.  
; REGISTRATION NUMBER: 31,815  
; REFERENCE/DOCKET NUMBER: P-CJ 1191  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (619) 535-9001  
; TELEFAX: (619) 535-8949  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2482 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-328-254-6

Query Match 3.7%; Score 104; DB 1; Length 2482;  
Best Local Similarity 21.4%; Pred. No. 1.6; 139; Indels 90; Gaps 17;  
Matches 80; Conservative 65; Mismatches 139; Indels 90; Gaps 17;  
QY 60 VHEHYD-CTLNQTNIGN-----NNKFIITQLLEGSRRFFCWNRWGVGVGQSKMNHFT 113  
Db 1825 LHNDOEACKAKEQNLSSQVECLEAKQLQGLDEAKNNVI-----VLQSSVNGLI 1875  
QY 114 C-LEDAKKDKPKKFE--KTKNKEERDRFVAOPNKYTLIEVOGEAE---SQEAVVKALS 167

Db 1876 QEVEDGQKLEKDEEISRLKNQIQDQQLVSK-----LSQVEGEHQLWKEQNLRLNT 1930  
QY 168 POYDSGVPVIVVPCSLDPATONLITNIFSKEMPKNMTLMNLDVKKMPLGKLTQOQIAR 227  
Db 1931 VELEQKIQVLOSNNASLQDTLEVLQS---SYKLNLELELTQMD--KMSFVYKNNKMTAK 1985  
QY 228 GFPALEALEEAMKNPTGDGQSLSEELSCFVTVPHPNFGSRPPPIPSDVLQAAKMDL-- 285  
Db 1986 ETELQREHME-MAQKTAELQ--EELSG-----EKNRLAGELQLLEEKSKDQLKE 2034  
QY 286 LVIADIELAOTLAAPGEEKEVEVPHPLDRDYQL-----LRCQLQLDSDGSEYKAI 339  
Db 2035 LTLSENSELKSLDCMHKQDQVEKGVREET-AEYQLRLHEAKKHQALLDNTNKQYEVEI 2093  
QY 340 QTYLKOTGNSYRC-----PMLRHVWKN 362  
Db 2094 QTYREKLTSKBECLSSQKLEIDLKSSKEELNNSLKATTOILEELKKTMDNLKYNOLK 2153  
QY 363 REGEDRFQAHSKL 376  
Db 2154 KENE--RAQGMKML 2165  
RESULT 6  
US-08-405-615-1  
; Sequence 1, Application US/08405615  
; Patent No. 5602095  
; GENERAL INFORMATION:  
; APPLICANT: Pastan, Ira  
; APPLICANT: Fitzgerald, David J.  
; TITLE OF INVENTION: Recombinant Pseudomonas Exotoxin with  
; TITLE OF INVENTION: Increased Activity  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Ellen L. Weber  
; STREET: One Market Plaza, Steuart Tower, Suite 2000  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94105  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/405,615  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/901,709  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 15280-36  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-543-9600  
; TELEFAX: 415-543-5043  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 613 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-405-615-1

Query Match 3.7%; Score 103.5; DB 1; Length 613;  
Best Local Similarity 19.7%; Pred. No. 0.19;  
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;



NAME: Weber, Ellen Lauver  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 15280-36-2  
TELEPHONE: (415) 576-0200  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 613 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-08-463-480-1

Query Match 3.7%; Score 103.5; DB 2; Length 613;  
Best Local Similarity 19.7%; Pred. No. 0.19;  
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;

QY 172 SGPVRTVVKP-----CSLDPATONLITNIFSKEMFNMTLMNDVK----- 213  
Db 192 SGKVLCLLDPLDGVNYLAQORCNLDDTWEGKIYRVLAGNPXKH-----DLIDKPTVISH 246  
QY 214 -----KMPGLKLTQOIARGFEALE----- 233  
Db 247 RLHPPEGSLAALTAHQACHLPLETFRHRQPRGWEQLEOCGYVQRLVALYLAARLSWN 306  
QY 234 -----ALEEAMKNTGDSLEELSSCFYTVIPHNFGRSPPINSPPVLQAKKMDLVLVA 289  
Db 307 QVDQVIRNALASP-GSGGDLGEAT-----REQPE-----QAR--LALTLA 343  
QY 290 DIELAQTLOAAPGEEKEVEVPHPLDRDYQLLRQQL-----LDSG-----ESEYKAI 339  
Db 344 AAESERFVROGTGDEAGA-----ANADVSLTCPVAAGECAGPADSGDALLERNPTG 397  
QY 340 QTYLKOTGNSYRCPNLRHVWKNREGGDRFQAHSKLGNNRLL---WHGTNVAVVAAILT 396  
Db 398 AEFLGCGDVFSFTRGTQNTVER-----LLQAHRLQLEERGYYVFGYHGTFLEAAQSIVF 452  
QY 397 SGLRIMPHSGGRVKGIGYFASSENSKAGYVTTMHCGHQVGYMFLGVALGKEHHITIDD 456  
Db 453 GGVARSQDLDAIWRGFYIAGDPALAYGAYAQDQEPDAR-----GRIRNGALLRVVYVPR 505  
QY 457 PSLSKSPPGDSVIARGOTEPDPAQDIELELDGQPV 492  
Db 506 SSL-----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 535

RESULT 9  
US-08-225-224-1  
Sequence 1, Application US/08225224  
Patent No. 5635599  
GENERAL INFORMATION:  
APPLICANT: PASTAN, Ira  
APPLICANT: KREITMAN, Robert J.  
TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND  
TITLE OF INVENTION: CIRCULARLY PERMUTATED FUSION PROTEINS  
NUMBER OF SEQUENCES: 57  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend Kourie and Crew  
STREET: Steuart Street Tower, One Market Plaza  
CITY: San Francisco  
STATE: California  
COUNTRY: US  
ZIP: 94105-1493  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/225,224  
FILING DATE: 8-APR-1994  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen L.  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 15280-193  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 614 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
FEATURE:  
NAME/KEY: Protein  
LOCATION: 1..614  
OTHER INFORMATION: /label= native-PE  
US-08-225-224-1

Query Match 3.7%; Score 103.5; DB 1; Length 614;  
Best Local Similarity 19.7%; Pred. No. 0.19;  
Matches 78; Conservative 47; Mismatches 144; Indels 127; Gaps 17;

QY 172 SGPVRTVVKP-----CSLDPATONLITNIFSKEMFNMTLMNDVK----- 213  
Db 193 SGKVLCLLDPLDGVNYLAQORCNLDDTWEGKIYRVLAGNPXKH-----DLIDKPTVISH 247  
QY 214 -----KMPGLKLTQOIARGFEALE----- 233  
Db 248 RLHPPEGSLAALTAHQACHLPLETFRHRQPRGWEQLEOCGYVQRLVALYLAARLSWN 307  
QY 234 -----ALEEAMKNTGDSLEELSSCFYTVIPHNFGRSPPINSPPVLQAKKMDLVLVA 289  
Db 308 QVDQVIRNALASP-GSGGDLGEAT-----REQPE-----QAR--LALTLA 344  
QY 290 DIELAQTLOAAPGEEKEVEVPHPLDRDYQLLRQQL-----LDSG-----ESEYKAI 339  
Db 345 AAESERFVROGTGDEAGA-----ANADVSLTCPVAAGECAGPADSGDALLERNPTG 398  
QY 340 QTYLKOTGNSYRCPNLRHVWKNREGGDRFQAHSKLGNNRLL---WHGTNVAVVAAILT 396  
Db 399 AEFLGCGDVFSFTRGTQNTVER-----LLQAHRLQLEERGYYVFGYHGTFLEAAQSIVF 453  
QY 397 SGLRIMPHSGGRVKGIGYFASSENSKAGYVTTMHCGHQVGYMFLGVALGKEHHITIDD 456  
Db 454 GGVARSQDLDAIWRGFYIAGDPALAYGAYAQDQEPDAR-----GRIRNGALLRVVYVPR 506  
QY 457 PSLSKSPPGDSVIARGOTEPDPAQDIELELDGQPV 492  
Db 507 SSL-----PGFYRT-SLTLAAPEAAGEVE-RLIGHPL 536

RESULT 10  
US-08-722-258-1  
Sequence 1, Application US/08722258  
Patent No. 6011002  
GENERAL INFORMATION:  
APPLICANT: PASTAN, Ira  
APPLICANT: KREITMAN, Robert J.  
APPLICANT: PURI, Raj K.  
TITLE OF INVENTION: Circularly Permuted Ligands and  
TITLE OF INVENTION: Circularly Permuted Chimeric Molecules  
NUMBER OF SEQUENCES: 72  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, Eighth Floor  
CITY: San Francisco  
STATE: California



Qy	172	SGPVRTVVKP-----C\$LDPAQNLTNTFSKEMFNAMTLMNLDVK-----	213
Db	217	SGKVCLLLDPLDGWYVLAQRORNLMDWEGKIYRVLAGNPAKH-----DLDIKPVTIVSH	271
Qy	214	-----KMPLGKLTQKOIANGFPALE-----	233
Db	272	RLHFPEGGSALAATAHQACHLPLETETRHRQPRGWEOEQCGYPQORLVALYAARLSWN	331
Qy	234	----ALEEAMKNPTGDQSILEELSCTFYVIPIHNFGRSRPPPIINSPDVLOAKDKMLLVIA	289
Db	332	OVDOVRNALASP-GSGGDLGEAI-----REQPE-----OAR--LALTIIA	368
Qy	290	DIELAOTLQAAPGEEKEVEVPHPILDRDYQLIRCOLQL-----LD\$G-----ESEYKAI	339
Db	369	AASERPVROGTND\$EAGA-----ANADVSLTC\$PAAGECAGPADSGDALLERNYPYG	422
Qy	340	QTYLKOTGNSYRCPNLRHWKVNREGEGRDFQAH\$KLGNRRLL--WHGTVNAVVAAILT	396
Db	423	A\$FLGDGGDV\$FSTRQTQNTVER-----LLQAHRLQEERG\$YVFVG\$YHCTFEAAQ\$IVF	477
Qy	397	\$GLRIMPHSGGRVKGKIYFAS\$EN\$KSAGIYVTMHCCHGVGMFTGEVALGHEHITIDD	456
Db	478	GGV\$ARSQDLDA\$WRGFYTAGDPALAYGAQOEOPDAR-----GRIRNGALLRRVYPR	530

[illegible]



```

: : : ||
Db 280 GDPALAYGYAQDQEPDAR-----GRIRNGALLRVVPRSSL-----PGFYRT-SLTLAA 327

: : : ||
Qy 477 PDPAODIELELDGQPV 492

: : : ||
Db 328 PEAGEVE-RLIGHPL 342

RESULT 14
US-08-839-425-7
; Sequence 7, Application US/08839425
; Patent No. 5912322
; GENERAL INFORMATION:
; APPLICANT: Riemen, Mark W
; APPLICANT: Stirdivant, Steven M
; TITLE OF INVENTION: Modified PE40
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Merck & Co., Inc.
; STREET: 126 Lincoln Avenue
; CITY: Rahway
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07065
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Power Macintosh 6.0.1
; SOFTWARE: Microsoftword 6.0.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/839,425
; FILING DATE:
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Grassler, Frank P
; REGISTRATION NUMBER: 31,164
; REFERENCE/DOCKET NUMBER: 17879IA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908) 594-3462
; TELEFAX: (908) 594-4720
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 420 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-839-425-7

Query Match 3.6%; Score 100.5; DB 2; Length 420;
Best Local Similarity 21.2%; Pred. No. 0.21;
Matches 67; Conservative 38; Mismatches 126; Indels 85; Gaps 14;

Qy 215 MPLGKLTQQIARGFEALE-----ALDEAMKNPTGDSGL 249
: : : ||
Db 74 LPLETLRHRQPRGWEQLEAGYVQRLVALYLAARLSWNQVQVIRNALASP-GSGDDL 132

Qy 250 EELSSCFYTVPHNFGSRPPPIINSPDVLQAKDMLVLADIQAQTLQAAPEEEKEVE 309
: : : ||
Db 133 GEAT-----REQPE-----QAR--LALTAAESESERFVROGTGNDGAGA- 169

Qy 310 EVPHPLDRDYOLLRCQLQL-----LDSG-----ESEYKAIQTYLKQGTNSVRCNLRHW 359
: : : ||
Db 170 -----ANADVSLTFCPVAAGSCAGPADSGDALLERNYPTAEFLGDDGDSVFSRGTQNW 224

Qy 360 KVNREGEEDRFQAHSKLGNRRL-----WHGTNVAVVAAILTSGLRIMPHSGRGGKIYFA 416
: : : ||
Db 225 TVER-----LQAHQRLQEEYGVVGVHGTFLQAQSVIFGCVRARQQDLDAIWRGYIA 279

Qy 417 SENSKSAGYVTHMCGHGQVGYMFLGVALGKEHHITDDPSLSPPPGFDVIRAGQTE 476
: : : ||
Db 280 GDPALAYGYAQDQEPDAR-----GRIRNGALLRVVPRSSL-----PGFYRT-SLTLAA 327
```

```

Qy 477 PDPAODIELELDGQPV 492

: : : ||
Db 328 PEAGEVE-RLIGHPL 342

RESULT 15
US-09-336-447A-15
; Sequence 15, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP41 AND USP42 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY.024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 15
; LENGTH: 889
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
; US-09-336-447A-15

Query Match 3.6%; Score 100.5; DB 4; Length 889;
Best Local Similarity 19.9%; Pred. No. 0.69;
Matches 101; Conservative 78; Mismatches 166; Indels 163; Gaps 25;

Qy 8 SVQTEGSKKQ-----RQGTEDDSFR-STAEALRAAPADNRVIRVDPSPFSPRNGIQ 59
: : : ||
Db 423 SVSGEKERQIVHVGAGRISNDSTDAVNGSOLYALAAAVDDNQY-----DIE 469

Qy 60 VHEDYDCTLNOTNIGNNNNKFIYIQLLEEGSRFFCWNWGR-----VGRVG 105
: : : ||
Db 470 KNQD-DIAKNQADIAKNQAD---IQTLNVGKELLNLSGLRIDQKADIDNNINHIYELA 525

Qy 106 QSKMNHFTCLDADAKDKFKKFWKTKNKEERDRFVAQPNKYTLIEVOGEAESOEAE---- 161
: : : ||
Db 526 QQQDQHSDDIKTLKNVEE-----GLLESLGLIDQKADLT 562

Qy 162 VVKALSPQVDSGVRVTVVVKPCSLDPATONLTNITFSKMFKNAMTLMMLDKMPLGLKLT 221
: : : ||
Db 563 DIKALESNVEEG-----LLDLSGLRIDQ--KADIAQNQANIQDL---AAYNELQ 606

Qy 222 KQIARGFEALEALEEAMKNPTGDSGLLEELSSCFYTVPHNFGSRPPPI----- 272
: : : ||
Db 607 DQYAKQTEAIDALNKASSENT---QNTIEDLAA--YNELQDAYAKQQTAEIDALNKASSE 661

Qy 273 NSPDVLQAKDMLVLADI-ELAQ-----TLQAAPGEEKEKVEVPHPLDRDYQ- 320
: : : ||
Db 562 NTQIANQADIANNNIYELAQQQDQHSDDIKTLAKASANTDRIAKNKADADASEFET 721

Qy 321 LLRCQLQLD-----SGSEYKAIQTYLKQGTNSVRCNLRHW 359
: : : ||
Db 722 LTKNQNTLIEKDEKHLITANKTAIDANKASADTKFAATADAITKNGNAI-TKNAKSIT 780

Qy 360 KVNREGE-----DRFOA-HSKLGNRRLWHGTNVAVVAAILTSGLRIM 402
: : : ||
Db 781 DLGTQVDFGDRVTALDTKVNAFDRITALDSKVEN-----GMAQAAL--SGL-FQ 829

Qy 403 PHSGRGGKIYFASENSKS-----AGY 425
: : : ||
Db 830 PYSVGKFNATAALGGYGSKSAVAGY 857

Search completed: August 29, 2002, 07:58:21
```

Job time: 298 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:52:23 ; Search time 124.84 seconds  
(without alignments)  
507.146 Million cell updates/sec

Title: US-09-701-586B-2  
Perfect score: 2998  
Sequence: 1 MAARRRRSTGGGRARALNES.....PNQVRMYRLKVKQNFELQLW 570

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11107396 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_032802.\*
- 1: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*
  - 2: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*
  - 3: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*
  - 4: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*
  - 5: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*
  - 6: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*
  - 7: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*
  - 8: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*
  - 9: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*
  - 10: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*
  - 11: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*
  - 12: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*
  - 13: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*
  - 14: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*
  - 15: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*
  - 16: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*
  - 17: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*
  - 18: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*
  - 19: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*
  - 20: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*
  - 21: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*
  - 22: /SIDSL1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2998	100.0	570	21	AAV51174 Human brain PARP2
2	2998	100.0	570	22	AAAB11480 Human brain poly-A
3	2981.5	99.4	583	22	AAAB47029 hPARP2. Homo sapi
4	2749	91.7	521	22	AAAB60693 Human poly(ADP-rib
5	2732.5	91.1	534	21	AAAB42909 Human ORFX ORF2673
6	2732.5	91.1	534	22	AAU29023 Human PARP-2 prote
7	2723.5	90.8	534	22	AAU29020 Human PARP-2 prote
8	2693.5	89.8	531	22	AAAB33513 Human protein sequ
9	2468	82.3	522	22	AAU29022 Mouse PARP-2 prote
10	2043.5	68.2	1063	22	AAAB47032 Fusion protein PAR
11	1966	65.6	379	22	AAU21687 Novel human neopla

12	1790.5	59.7	360	22	AAAB47030 N-terminal fragmen
13	1518	50.6	287	22	AAAB47031 C-terminal fragmen
14	1456	48.6	294	22	AAU20129 Human DNA repair a
15	1456	48.6	294	22	AAU21810 Novel human neopla
16	1180	39.4	637	21	AAAY68835 The poly(ADP-ribos
17	1149.5	38.3	653	21	AAAY68834 A poly(ADP-ribose)
18	1145	38.2	227	22	AAU20130 Human DNA repair a
19	1145	38.2	227	22	AAU21811 Novel human neopla
20	1012.5	33.8	1013	17	AAAR99642 Poly(ADP-ribose) p
21	1010.5	33.7	1014	20	AAAY33699 Human poly(ADP-ri
22	1007.5	33.6	1014	21	AAAY58043 Human PARP-1 prote
23	1007.5	33.6	1014	22	AAU29019 Human tankyrase2 r
24	1007.5	33.6	1014	22	AAAB66296 Human nuclear NAD+
25	998.5	33.3	1014	21	AAAY49939 Drosophila melanog
26	979.5	32.7	557	22	ABB66431 A poly(ADP-ribose)
27	961.5	32.1	969	21	AAAY68833 A poly(ADP-ribose)
28	961.5	32.1	980	21	AAAY68839 A poly(ADP-ribose)
29	911	30.4	190	22	AAU21688 Maize poly ADP-rib
30	847.5	28.3	982	20	AAAY28464 Human brain PARP3
31	725	24.2	533	21	AAAY51175 Human uterus type
32	725	24.2	540	21	AAAY51176 Human PARP-3 prote
33	720	24.0	533	22	AAU29021 Murine PARP1 (shor
34	700.5	23.4	528	21	AAAY51178 Murine PARP1 (long
35	698	23.3	533	21	AAAY51177 Human secreted pro
36	637	21.2	122	21	AAAG01000 Fusion protein of
37	576	19.2	1010	21	AAAY68840 cDNA sequence enco
38	306.5	10.2	1724	21	AAAY54373 Human minor vault
39	306.5	10.2	1724	22	AAAB51022 Human tankyrase2 e
40	304.5	10.2	1730	22	AAU33242 Mouse SPANK. Mus
41	299	10.0	1099	22	AAAB66301 Novel human secreto
42	141	4.7	991	22	AAAB47023 Novel human diagno
43	139.5	4.7	1515	22	ABG10624 Novel human diagno
44	136.5	4.6	1442	22	ABG25601 Novel human diagno
45	136.5	4.6	2408	22	ABG10631

ALIGNMENTS

RESULT 1  
ID AAY51174 standard; Protein: 570 AA.  
AC AAY51174;  
DT 31-MAR-2000 (first entry)  
XX Human brain PARP2 protein.

DE PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage.

OS Homo sapiens.  
XX WO9964572-A2.  
XX 16-DEC-1999.

XX 04-JUN-1999; 99WO-EP03889.  
XX 05-JUN-1998; 98DE-1025213.  
XX 01-MAR-1999; 99DE-1008837.  
(BADI ) BASF AG.

XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX WPI: 2000-087218/07.  
XX N-PSDB; AAZ44287.

XX Novel genes and proteins, antibodies and binding partners useful in  
XX diagnosis and therapy of energy deficiency associated disease

PT conditions -  
PS Claim 4; Page 52-54; 96pp; German.  
XX  
CC This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>2MHX<sub>2</sub>2C (1). The nucleic acid  
CC sequences, PARP homologues and antibodies are useful for analytic  
CC detection of PARP homologues and for identifying PARP effectors or  
CC binding partners, as well as for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
CC human PARP2 protein used in the method of the invention.  
XX  
SQ Sequence 570 AA;  
  
Query Match 100.0%; Score 2998; DB 21; Length 570;  
Best Local Similarity 100.0%; Pred. No. 2.6e-275;  
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 MAARRRSTGGRRARALNESKRVNNGTAPEDSSPAKTRRCQROESKMPVAGGKANKD 60  
DB 1 maarrstggrraralneskrvnnngtapedsspaktrrcrqroeskmvpaggkankd 60  
  
QY 61 RTEKDQESVKALLKLGKAPVDPCTAKGVKAHYVCGNDVYDMLNQTQFNKKYYL 120  
DB 61 rtekdqesvkallkqapvdpctakgvkayhvcgndvymlnqtnlqfnnkyyll 120  
  
QY 121 IQLLEDDAQRNFSVMMRWGRVGMQSHLVACSGNLNKAKEIFQKFLDKTKNNWEDREK 180  
DB 121 iqlleddaqrnfsvmmrwgrvgmqshlvacsgnlmkakeifqkfldtktnnwedrek 180  
  
QY 181 FEKVPQGYDMLQNDYANTODEETKEESLKSPLKPESOLDLVRQELIKLCNVQAMEE 240  
DB 181 fekvpgkydmlqndyantodeetkeeslksplkpesoldlrvqeliklcnvqamee 240  
  
QY 241 MMEEMKYNTKKAPLGKLTVAQIRAGYQSLKKIEDICIRAGQGRALMEACNEFYTRIPHDF 300  
DB 241 mmemkynntkkaplgkltvaqiragyslkkiediciragqgralmeacnefytriphdf 300  
  
QY 301 GLRTPPLIRTKQELSEKIQLEALGDIEIAIKLVKTELOQSPHPLDOHYRNLCALRPLD 360  
DB 301 glrtpplirtkqelsekiqulealgdieiaiklvkteleqsphepldohyrnlcalrpld 360  
  
QY 361 HESYEFKVISQYLSQTHAPTHSDYTMFLDLDFEVEKDGKEAFREDLHNRMLLWHGSRMS 420  
DB 361 hesyefkvisqylsqthapthsdymtflldlfevekdgekeafredlhnmllwhgsrms 420  
  
QY 421 NWGILSHGLRIAPPEAPITGYMFGKGIYFADMSKKSANYCFASRLKNTGLLLSEVALG 480  
DB 421 nwgilshglriappeapitgymfgkgiyfadmskksanycfarslknrtglllsevalg 480  
  
QY 481 QCNELLEANPKAEGLLQGSTKGLGRKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYT 540  
DB 481 qcnelleanpkagellqgstkgkgrkmapssahfvtlngstvpplgpasdtgilnpgdyt 540  
  
QY 541 LNYNEYIVYNPQVNRMYLLKLVQFNFLQLW 570  
DB 541 lnyneyivynpqvnmryllkvqfnflqlw 570  
  
RESULT 2  
AAB11480  
ID AAB11480 standard; Protein; 570 AA.  
XX  
AC AAB11480;  
XX

DT 02-MAR-2001 (first entry)  
XX  
DE Human brain poly-ADP-ribose-polymerase protein.  
XX  
KW Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;  
KW cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;  
KW anticonvulsant; cytosatic; antibacterial; immunosuppressive; treatment;  
KW antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;  
KW 2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;  
KW neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;  
KW Parkinson's disease; ischemic damage; microinfarction; sepsis;  
XX diabetes mellitus.  
OS Homo sapiens.  
XX  
XX DB19921567-Al.  
XX  
XX 16-NOV-2000.  
XX  
XX 11-MAY-1999; 99DE-1021567.  
XX  
XX 11-MAY-1999; 99DE-1021567.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Lubisch W, Sadowski J, Kock M, Hoeger T;  
PI  
XX WPI; 2001-032983/05.  
XX  
XX N-PSDB; AAC82090.  
XX  
XX Drugs for inhibiting PARP or especially homologous enzymes comprising  
XX 4-substituted phthalazinone derivatives, useful e.g. for treating  
XX neurodegenerative disease, ischemic damage, tumors or diabetes -  
XX  
XX Example A; Page 12-13; 14pp; German.  
XX  
XX This invention describes novel 4-substituted 2H-phthalazin-1-one  
XX derivatives (I) which are used for the treatment or prophylaxis of  
XX diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;  
XX EC 2.4.2.30) activity. The products of the invention have nootropic,  
XX neuroprotective, cerebroprotective, antiparkinsonian, nephrotropic,  
XX cardiant, vasotropic, anticonvulsant, cytostatic, antibacterial,  
XX immunosuppressive, antiinflammatory, antirheumatic, antiarthritic,  
XX antidiabetic. (I) are especially used for treating or preventing  
XX neurodegenerative disease or neuronal damage (specifically associated  
XX with ischemia, trauma or massive bleeding, especially apoplexy or  
XX spinal-cranial trauma; or Alzheimer's disease, Huntington's disease or  
XX Parkinson's disease), treating or preventing ischemic damage  
XX (specifically renal damage after renal ischemia or during and after  
XX kidney transplantation or heart damage after cardiac ischemia), treating  
XX epilepsy, specifically generalized epileptic attacks (e.g. petit mal and  
XX tonic-clonic attacks) or partial epileptic attacks (e.g. temporal lobe  
XX and complex partial attacks), treating microinfarction (e.g. during and  
XX after heart valve replacement, aneurysm resectioning and heart  
XX transplantation), revascularization of critically constricted coronary  
XX arteries (e.g. after PCTA or by-pass operations) or peripheral arteries  
XX (e.g. leg arteries), treating acute myocardial ischemia and damage during  
XX or after its mechanical or drug-induced lysis and treating tumors and  
XX their metastasis, sepsis and septic shock, inflammatory and rheumatic  
XX disease (e.g. rheumatoid arthritis), and diabetes mellitus. Although (I)  
XX inhibit PARP (i.e. the known form designated PARP1), they especially  
XX selectively and strongly inhibit PARP homologs, specifically the homolog  
XX PARP2. In particular (I) have very strong PARP2 inhibitory activity (e.g.  
XX with K<sub>i</sub> values of 1-20 nM) and high selectivity for PARP2 relative to  
XX PARP1 (generally by a factor of more than 5).  
XX  
SQ Sequence 570 AA;  
  
Query Match 100.0%; Score 2998; DB 22; Length 570;  
Best Local Similarity 100.0%; Pred. No. 2.6e-275;  
Matches 570; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS	Homo sapiens.	
XX		
PN	WO200077179-A2.	
XX		
PD	21-DEC-2000.	
XX		
PF	16-JUN-2000; 2000WO-US16629.	
XX		
PR	16-JUN-1999; 99US-0139543.	
PR		
PA	(ICOS-) ICOS CORP.	
XX		
PI	Christenson E, Demaggio AJ, Goldman PS, McElligott DL;	
DR		
XX		
DR	WPI; 2001-0253335/03.	
DR	N-PSDB; AAC85303.	
XX		
PT	New human poly(ADP-ribose) polymerase for treating inflammatory,	
PT	neurological, cardiovascular, or neoplastic tissue growth disorders,	
PT	such as, arthritis, encephalitis, myocardial ischemia, and leukocyte	
PT	metastasis -	
XX		
XX		
PS	Claim. 3; Page 94-95; 129pp; English.	
XX		
CC	This sequence represents human poly(ADP-ribose) polymerase (hPARP2).	
CC	This protein causes the covalent addition of polymers of ADP-ribose	
CC	to protein targets. hPARP2 activity is induced in many instances of	
CC	oxidative stress or during inflammation where there is direct damage	
CC	to the DNA. hPARP2 may be used to identify antagonists which	
CC	may be used to treat a human having a disorder mediated by PARP2	
CC	activity, such as, inflammatory, neurological, cardiovascular,	
CC	or neoplastic tissue growth disorders, e.g. ischemic stroke,	
CC	hemorrhagic shock, myocardial ischemia or infarction,	
CC	transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty	
CC	arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic	
CC	shock, gram negative or positive sepsis, toxic shock syndrome; multiple	
CC	organ injury syndrome secondary to septicemia, trauma, or hemorrhage;	
CC	allergic or vernal conjunctivitis, uveitis, thyroid-associated	
CC	ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis,	

CC sarcosis, pulmonary sarcoidosis, pneumonia, abscessitis, vasculitis, vasculitides,  
CC sarcosis, bronchiectasis, pulmonary oxygen toxicity, reperfusion  
CC injury of the myocardium, brain or extremities; cystic fibrosis; keloid  
CC formation, scar tissue formation; atherosclerosis; systemic lupus  
CC erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's  
CC syndrome; graft versus host disease, allograft rejection; chronic  
CC glomerulonephritis; inflammatory bowel disease, Crohn's disease,  
CC ulcerative colitis, necrotizing enterocolitis; inflammatory dermatoses  
CC contact or atopic dermatitis, psoriasis, urticaria, fever and myalgia  
CC due to infection: meningitis, encephalitis, and brain and spinal cord

[illegible]

QY 108 QTNLFQNNKYLIQLLEDAAQRNFSYVMRWGRVYGMQHSVLVACSGNLNKAKEIFOKKF 167  
DB 121 qtnlfqnnkyyliqlledaaqrnfsvmwrwgrvymqgshslvacsgnlkakakeifqkf 180  
QY 168 LDKTKNWEDEKFEKVPKGYDMLQNDYATNTQDEEETKKEESLKSPLKPSQDLRLVQE 227  
DB 181 ldktknweдекфквпкгьдмлқндьатнтқдеетккееслксплкпсқдлрлвqe 240  
QY 228 LKLCNVQAMEEMMEMKYNKKAPLGKLTVAQIKAGYOSLKIEDCIRAGQHGRALME 287  
DB 241 lklicnvqameemmemkynkkaplgkltvaqikagyslkiekciragqhggralme 300  
QY 288 ACNEFYTRIPDFGLRPPPLRTOKELSEKLTOLLEALGDIEATKLVKTELQSPHEPLDQ 347  
DB 301 acnefytriphdfglrppplrtokelsekltolleaigdieatklvktelqsphepldq 360  
QY 348 HYRLHLCALPLDHSEYEFKVISYQLOSTHAPTHSDYTMLLDLFEYKDGKEAFREDL 407  
DB 361 hyrlhcalpldhsefefkvisyqlqsthapthsdymtlldlfevekdgkeafredl 420  
QY 408 HNRMLLHGSRMSNWVGLSHGLRIAPPEAPITGYMFGKGIYFADMSKSNYCFASRLK 467  
DB 421 hnrmlhwgsrmsnwvglshglriappeapitgymfgkgyifadmsksanycfasrlk 480  
QY 468 NTGLLSEVALGQCNEELLEANPKAEGILQKHSKTKGLGKMAPSSAHFVTLNGSTVPLGP 527  
DB 481 ntgllsevalgcneelleanpkaelqgkhsstkglgkmapssahfvtlngstvpplgp 540  
QY 528 ASDTGILNPDGYTLNNEYIVYNPNQVMRYLLKVFNFQLQW 570  
DB 541 asdtgilnpgdgytlneyivynpnqvmryllkvqnfqlqw 583

## RESULT 4

AAB60693  
ID AAB60693 standard; Protein; 521 AA.

XX AC AAB60693;

XX DT 11-MAY-2001 (first entry)

XX DE Human poly(ADP-ribose) synthetase sbhPARS2.

XX KW Human; poly(ADP-ribose) synthetase; sbhPARS2; vaccine; drug screening;  
KW ischaemic disorder; cerebral ischaemia; cardiac ischaemia;  
KW myocardial infarction; stroke; inflammation; autoimmune disease;  
KW diabetes; multiple sclerosis; neurodegenerative disease;  
KW Parkinson's disease; Alzheimer's disease; chromosome localisation.

XX OS Homo sapiens.

XX PN W0200112645-A1.

XX PD 22-FEB-2001.

XX PF 10-AUG-2000; 2000WO-US21775.

XX PR 12-AUG-1999; 99US-0373441.

XX PA (SMIK ) SMITHKLINE BEECHAM CORP.

XX PI Barone F, Field J, Kabnick KS, Li X, McQueney MS, Zhu Y;

XX DR WPI; 2001-2111196/21.

XX DR N-PSDB; AAF59996.

XX PT New human polypeptide of the polyADPribose synthetase family for use in  
XX diagnostic assays and for screening modulators used for preventing and  
XX treating inflammation, autoimmune disease and Alzheimer's disease

XX PS Claim 1; Page 21; 30pp; English.

XX XX

CC This sequence represents a novel human poly(ADP-ribose)  
CC synthetase, sbhPARS2. The invention also relates to fragments,  
CC variants and sequences with at least 95% identity to the sbhPARS2  
CC protein or nucleotide sequence; expression systems and host cells  
CC comprising an sbhPARS2 nucleic acid sequence; the recombinant expression  
CC of sbhPARS2; and an antibody specific for sbhPARS2. sbhPARS2 proteins  
CC and nucleotides are useful as vaccines for inducing an immunological  
CC response in a mammal. The sbhPARS2 protein is useful for identifying  
CC compounds which inhibit or stimulate its activity or expression level.  
CC Such agonists and antagonists of sbhPARS2 are useful for treating human  
CC diseases including ischaemia and ischaemic tissue injury (e.g., cerebral  
CC and cardiac ischaemia, myocardial infarction, stroke), inflammation,  
CC autoimmune disease (e.g. diabetes, multiple sclerosis) and  
CC neurodegenerative diseases (e.g., Parkinson's disease and Alzheimer's  
CC disease). sbhPARS2 nucleic acids are useful as diagnostic reagents for  
CC detecting mutations in the associated gene; as hybridisation probes for  
CC to isolate full-length sbhPARS2 cDNAs and sbhPARS2 genomic clones; and  
CC for chromosome localisation studies. The sbhPARS2 protein is also useful  
CC as an immunogen to produce antibodies for therapeutic use. sbhPARS2  
CC proteins, nucleotides and antibodies are also useful in screening methods  
CC for detecting the effect of added compounds on the production of mRNA and  
CC protein in cells.

XX XX Sequence 521 AA;

Query Match 91.7%; Score 2749; DB 22; Length 521;  
Best Local Similarity 100.0%; Pred. No. 9.9e-252;  
Matches 521; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 50 MPVAGGKANKDRTEDEKQDESVKALLKKGKAPVDPPECTAKYGAHVYCEGNDVYDMLNQT 109

DB 1 mpvaggkankdrtekdqdesvkallkgkavpdppectakvgkavhvycegnvdyvdmnqt 60

QY 110 NLQFNNNKYLIQLLEDAAQRNFSVWRWGRVYGMQHSVLVACSGNLNKAKEIFQKFLD 169

DB 61 nlqfnnnkyyliqlledaaqrnfsvmwrwgrvymqgshslvacsgnlkakeifqkfld 120

QY 170 KTKNNWEDREKFEKVPKGYDMLQNDYATNTQDEEETKKEESLKSPLKPSQDLRLVQELI 229

DB 121 ktknnwedrekfekevpykgydmlqndyatntqdeetkkeeslkpsqdlrlvqeli 180

QY 230 KLICNVQAMEEMMEMKYNKKAPLGKLTVAQIKAGYOSLKIEDCIRAGQHGRALMEAC 289

DB 181 klcnvqameemmemkynkkaplgkltvaqikagyslkiekciragqhggralmeac 240

QY 290 NEFYTRIPDFGLRTPPLRTOKELSEKIQLEALGDIEATKLVKTELQSPHEPLDQHY 349

DB 241 nefytriphdfglrtpplirtokelseqillealgdieatklvktelqspehpldqhy 300

QY 350 RNHLHLCALPLDHSEYEFKVISYQLOSTHAPTHSDYTMLLDLFEVEKDGKEAFREDLHN 409

DB 301 rnhlcalpldhsefefkvisyqlqsthapthsdymtlldlfevekdgkeafredlhn 360

QY 410 RMLLHWSGRMSNWVGLSHGLRIAPPEAPITGYMFGKGIYFADMSKSNYCFASRLKNT 469

DB 361 rmlllhwgrmsnwvglshglriappeapitgymfgkgyifadmsksanycfasrlknt 420

QY 470 GLLLSEVALGQCNEELLEANPKAEGILQKHSKTKGLGKMAPSSAHFVTLNGSTVPLGPAS 529

DB 421 glillsevalgcneelleanpkaelqgkhsstkglgkmapssahfvtlngstvpplgpas 480

QY 530 DTGILNPDGYTLNNEYIVYNPNQVMRYLLKVFNFQLQW 570

DB 481 dtgilnpgdgytlneyivynpnqvmryllkvqnfqlqw 521

RESULT 5

AAB42909

ID AAB42909 standard; Protein; 534 AA.

XX XX

AC AAB42909;

XX XX

08-FEB-2001 (first entry)  
Human ORFX ORF2673 polypeptide sequence SEQ ID NO:5346.  
Human: open reading frame; ORFX: detection; cytostatic; hepatotropic;  
vulnerary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
hypotensive; dermatological; immunosuppressive; antiinflammatory;  
antiviral; antibacterial; antifungal; antirheumatic; antichyroid;  
neuroaemic; gene therapy; cancer; proliferative disorder; hypertension;  
neurodegenerative disorder; osteoarthritis; graft vs host disease;  
cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;  
cholesterol ester storage; systemic lupus erythematosus; infection;  
severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
bone damage; cartilage damage; antiinflammatory disease; coagulation;  
thrombosis; contraceptive.  
Homo sapiens.  
WO200058473-A2.  
05-OCT-2000.  
31-MAR-2000; 2000WO-US08621.  
31-MAR-1999; 99US-0127607.  
02-APR-1999; 99US-0127636.  
05-APR-1999; 99US-0127728.  
30-MAR-2000; 2000US-0540763.  
(CURA-) CURAGEN CORP.  
Shinkets RA, Leach M;  
WPI; 2000-602362/57.  
N-PSDB; AAC77118.  
Novel nucleic acids and peptides derived from open reading frame X,  
useful for treating e.g. cancers, proliferative disorders,  
neurodegenerative disorders and cardiovascular disease -  
Claim 11; Page 4522-4524; 5507pp; English.  
AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
which represent the human ORFX open reading frames 1 to 3161. The ORFX  
sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
antidiabetic; hypotensive; dermatological; immunosuppressive;  
antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
antichyroid; and antianaemic. The sequences can be used for determining  
the presence of or predisposition to, or preventing or treating  
pathological conditions associated with an ORFX-associated disorder. The  
nucleic acids can be used to express ORFX proteins in gene therapy  
vectors. The proteins and nucleic acids may be used to treat cancers,  
proliferative disorders, neurodegenerative disorders, osteoarthritis,  
graft vs host disease, cardiovascular disease, diabetes mellitus,  
hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
coagulation; to inhibit thrombosis; and as a contraceptive.  
Sequence 534 AA;

Query Match 91.1%; Score 2732.5; DB 21; Length 534;  
Best Local Similarity 97.6%; Pred. No. 3.8e-250;  
Matches 521; Conservative 0; Mismatches 0; Indels 13; Gaps 1;

QY	50	MPVAGGKANKDRTEKDQD-----ESVKALLLKGKAPVDPECTAKYGAHVYC	96
Db	1	mpvaggkankdrtdedkqgmpgrswaskrvsesvskallllkgkpvdpsectakvgkahvyc	60
QY	97	EGNDVYDVMNLNQTNPNKNNKYYLIOLLEDDAQRFNSVMWRGVRGKMGHSLVACSGNL	156
Db	61	egndvdyvmnlqnlqfnnnkyyliqlleddaqrfnsvvmrgrvrgkmgqhsivacsgnl	120
QY	157	NKAKEFQKFKFDKTKNNWEDREKFEKVPCKYDMLQMDYATNTQDEETKEESLSPLK	216
Db	121	nkakeifqkfkfdtknnwedrekfekvpqkydmlqmdyatntqdeetkkeslsplk	180
QY	217	PSQLDLRVQELIKLTCNVQAMEEMEMKYNTPKAPLGLTVAQIKAGYQSLKIKIEDCI	276
Db	181	psqldlrvqeliklcnvqameememkynntkklpklvtvaqikagyslkkiedci	240
QY	277	RAGQHGRLAMEACNEFYTRIPHDGFLRTPLIRTKELSEKIQLLLEALGSDIEIAIKLVKT	336
Db	241	ragqhgrralmeacnefytriphdfgrtpplirtkelskqiqllealgieiaiklvkt	300
QY	337	ELQSPHPLDQHYRNLCALRPLDHESYEFKVISQYLOSTHAPTHSDYNTMLDLFEVEK	396
Db	301	elqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdymtllldlfevek	360
QY	397	DGEKEAFREDLHNRMLLWHGSRMSNVGILSHGLRIAPPEAITGYMFGKGIYFADMSSK	456
Db	361	dgekeafredlhnrmllwhgsrmsnvvgllshglrlrtappeaitgymfkgiyfadmssk	420
QY	457	SANYCFASRLKNTGLLLSEVALGQCNELEANPKAEGLLQGRKSTKGLGKMAPSSAHFY	516
Db	421	sanycfasrlkntglllsevalgqcneleankpaelgllqgkstkglgkmapssahfy	480
QY	517	TLNGSTVPILGPASDTGILNPDGYTLNXYINYNQVPMRYLLKVFQNFLOLW	570
Db	481	tlngstvpilgpasdtglnpdgytlnyeyivynpqvmryllkvqnfqlqiw	534
RESULT 6			
AAU29023			
ID	AAU29023	standard; Protein; 534 AA.	
XX	AAU29023;		
XX	18-DEC-2001	(first entry)	
DT	XX	Human PARP-2 protein #2.	
DE	XX		
KW	Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;		
KW	cytostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;		
KW	immunosuppressant; hyperproliferative disorder; cancer; cellular injury;		
KW	oxidative stress; neuroproliferative disorder; parkinsonism; apoptosis;		
KW	meningitis-associated intracranial complication; ischaemia;		
KW	inflammatory disorder; autoimmune disorder; arthritis; diabetes.		
OS	Homo sapiens.		
XX	WO200164955-A1.		
PN	07-SEP-2001.		
PD	01-MAR-2001; 2001WO-US06572.		
PF	02-MAR-2000; 2000US-0517467.		
XX	(ISIS-) ISIS PHARM INC.		
PI	Popoff I, Cowser LM;		
XX	WPI; 2001-602570/68.		
DR	N-PSDB; AAS45684.		
XX	Antisense compound useful for treating hyperproliferative,		

PT neurological, inflammatory and autoimmune disorders and diabetes  
PT inhibits human PARP -  
PS Example 16; Page 125-127; 168pp; English.  
XX The invention relates to antisense oligonucleotides targeted to human  
CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, DNA repair, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARP in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARP especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
CC arthritis) and diabetes. The present sequence is a PARP protein,  
CC the cDNA encoding which was used to design the antisense  
XX oligonucleotides.  
XX  
SQ Sequence 534 AA;

Query Match 91.1%; Score 2732.5; DB 22; Length 534;  
Best Local Similarity 97.6%; Pred. No. 3.8e-250;  
Matches 521; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
QY 50 MPVAGGKANKORTEDKQD-----ESVKALLLKGKAPVDPECTAKVGAHVYC 96  
Db 1 mpvaggkankdrtdkqdgmpgrswskrvsvskallllkgkapydpectakvgaahvc 60  
QY 97 EGNVDYVLMNQTNIQFNNNKYYLIQLLEDQAQRNFSVMMRWGRVKGQHSVLVACSGNL 156  
Db 61 egndvydvmlnqtnlqfnnnkyyllqllledaqrnfsvmmrwgrvkgngqhsvlvacsgnl 120  
QY 157 NKAKEIFOKKFLDKTKNNWEDREKFEKVPKGKDYMLQMDYATNTQDEETKKEESLKSPLK 216  
Db 121 nkakeifgkfkldkcknnwedrekfekvpgkydmlqmdyatntqdeetkkeslksplk 180  
QY 217 PESQLDLRVQELIKLICNVQAEEMMEMKYNTRKAPLGKLTVAQIKAGYQSLKKIEDCI 276  
Db 181 pesqldlrqvqeliklicnvqaeemmemkyntrkkgltvqikagygslkkiedci 240  
QY 277 RAGQHGRALMEACNEFYTRIPHDGFLRTPPLIRTPQKELSEKIQLEALGDIEAIKLVKT 336  
Db 241 ragqhgralmeacnefytriphdfglrtppllrtqkelskqlilealgdieaialkvkt 300  
QY 337 ELQSPHEPLDQHYRNLCALRPLDHESYEFKVISYLOSTHAPTHSDYTMTLDDLFEVEK 396  
Db 301 elqsphepldqhyrnlcalrpldhesyefkvisyqlsthapthsdymtllldlfevek 360  
QY 397 DGEKEAFREDLHNRMLWHGSRMSNNVGLSHGLRIAPPEAPITCYMGFKGTYFADMSK 456  
Db 361 dgekeafredlhnrmllwhgsrmsnnvglshgllriappeapitcymfgktyfadmsk 420  
QY 457 SANYCFASRLKNTGLLLSEVALGQCNELEAPNKAEGLLQCKHSTKGLGKMAPSAHFV 516  
Db 421 sanycfasrlkntglllsevalgqcneleapnkaegllqckhstkglgkmapsaahfv 480  
QY 517 TLNGSTVPLGPASDFGIINPDGTYTLNNEYIVYNPNOVMRYLLKVQNFQLQLW 570  
Db 481 tlngstvpplgsdgtgiinpdgtytlnneyivynpnovmryllkvqnfqlqlw 534

## RESULT 7

AAU29020  
ID AAU29020 standard; Protein; 534 AA.

XX  
AC AAU29020;

XX  
DT 18-DEC-2001 (first entry)

XX  
DE Human PARP-2 protein #1.

XX Human: PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
KW cystostatic; nootropic; neuroprotective; antiinflammatory; antidiabetic;  
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
XX Homo sapiens.  
XX WO2000164955-A1.  
PN 07-SEP-2001.  
XX 01-MAR-2001; 2001WO-US06572.  
XX 02-MAR-2000; 2000US-0517467.  
PR (ISIS-) ISIS PHARM INC.  
XX  
XX Popoff I, Cowser LM;  
XX  
XX WPI; 2001-602570/68.  
DR N-PSDB; AAS45586.  
XX  
PT Antisense compound useful for treating hyperproliferative,  
PT neurological, inflammatory and autoimmune disorders and diabetes  
PT inhibits human PARP -  
XX  
PS Example 13; Page 102-104; 168pp; English.  
XX The invention relates to antisense oligonucleotides targeted to human  
CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARP in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARP especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
CC arthritis) and diabetes. The present sequence is a PARP protein,  
CC the cDNA encoding which was used to design the antisense  
XX oligonucleotides.  
XX  
SQ Sequence 534 AA;

Query Match 90.8%; Score 2723.5; DB 22; Length 534;  
Best Local Similarity 97.4%; Pred. No. 2.7e-249;  
Matches 520; Conservative 0; Mismatches 1; Indels 13; Gaps 1;  
QY 50 MPVAGGKANKORTEDKQD-----ESVKALLLKGKAPVDPECTAKVGAHVYC 96  
Db 1 mpvaggkankdrtdkqdgmpgrswskrvsvskallllkgkapydpectakvgaahvc 60  
QY 97 EGNVDYVLMNQTNIQFNNNKYYLIQLLEDQAQRNFSVMMRWGRVKGQHSVLVACSGNL 156  
Db 61 egndvydvmlnqtnlqfnnnkyyllqllledaqrnfsvmmrwgrvkgngqhsvlvacsgnl 120  
QY 157 NKAKEIFOKKFLDKTKNNWEDREKFEKVPKGKDYMLQMDYATNTQDEETKKEESLKSPLK 216  
Db 121 nkakeifgkfkldkcknnwedrekfekvpgkydmlqmdyatntqdeetkkeslksplk 180  
QY 217 PESQLDLRVQELIKLICNVQAEEMMEMKYNTRKAPLGKLTVAQIKAGYQSLKKIEDCI 276  
Db 181 pesqldlrqvqeliklicnvqaeemmemkyntrkkgltvqikagygslkkiedci 240  
QY 277 RAGQHGRALMEACNEFYTRIPHDGFLRTPPLIRTPQKELSEKIQLEALGDIEAIKLVKT 336  
Db 241 ragqhgralmeacnefytriphdfglrtppllrtqkelskqlilealgdieaialkvkt 300



QY 337 ELQSPHPLDQHYRNHLCALRPLDHSYEFKVISQYLOSTHAPTHSDYTMTLDLDFEVK 396  
|||||  
Db 301 elqsphepldqhyrnlhcalrpldhsyefkvisyqlsthapthsdymtlldlfevek 360  
QY 397 DGEKEAFREDLHNRMLLWHGSRMSNMVGLSHGLRIAPPEAPITGYMFGKGIYFADMSK 456  
|||||  
Db 361 dgekeafredlhnrmllwhgsrmsnmvlgshglriahpeapitgymfgkgyfadmsk 420  
QY 457 SANYCFASRLKWTGTLLELSEVALGQCNELEAPRAEGLLGKHSKGLGKMAPSSAHFV 516  
|||||  
Db 421 sanycfasrlkwtgltllesevalgqcnelleapraeagllgkhsckglgkmapssahfv 480  
QY 517 TLNGSVPLGPDASDTGILNPDYGTLYNNEYIVYNPNQVIMRYLLKVQFNFLQW 570  
|||||  
Db 481 tlngstvpilgpdasdtgilnpdygtlynyeyivynpnqvimryllkvqfnflqlw 534

RESULT 8  
ID AAB93513 standard; Protein; 531 AA.  
XX AAB93513;  
XX  
XX 26-JUN-2001 (first entry)  
XX Human protein sequence SEQ ID NO:12844.  
XX  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
XX Homo sapiens.  
XX EP1074617-A2.  
XX  
XX 07-FEB-2001.  
XX  
XX 28-JUL-2000; 2000EP-0116126.  
XX  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs -  
XX  
XX Claim 8; SEQ ID 12844; 2537pp + CD ROM; English.  
XX  
XX The present invention describes primer sets for synthesising 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dT primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesising polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by

CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. RAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to  
CC AAB95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
SQ Sequence 531 AA;  
  
Query Match 89.8%; Score 2693.5; DB 22; Length 531;  
Best Local Similarity 97.4%; Pred. No. 1.9e-246;  
Matches 517; Conservative 1; Mismatches 0; Indels 13; Gaps 1;  
  
QY 1 MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKTRRCORQESKMPVAGGKANKD 60  
|||||  
Db 1 maarrstgggraralneskrvnngntapedsspaktrrcqrqeskkmpvaggkankd 60  
QY 61 RTEDKQD-----ESVKALLKGAPVDPDECTAKVKGKHAHVYCEGNDVYDVLN 107  
|||||  
Db 61 rtedkqdgmpgrswaskrvsesvkalllkgapvdpectakvghvycgndvydvlm 120  
QY 108 QTNLOFNNNKYLLIQLLEDDAQRFNSVMRWGRVGMQHSLVACSGNLNKAKEIFQKPF 167  
|||||  
Db 121 qtnlqfnnnkyyliqlleddagrnfsvmrgvrgmgqhslvacsgnlnkakeifqk 180  
QY 168 LDKTKNNWEDREKFEKVPCKYDMLQNDYATNTQDEETKEESLSKPLPESQLDLRVQE 227  
|||||  
Db 181 ldktknnwedrekfekvpckydmlqndyatntqdeetkkeslskplkpesqldlrvqe 240  
QY 228 LKLCINVOAMEEMMEMKYNTKKAPLGLITVAQIKAGYQSLKKIEDCIRAGQHGRLME 287  
|||||  
Db 241 lklicnvqameemmemkynntkkaplgltvraqikagyslkkiedcragqhgrralme 300  
QY 288 ACNEFYTRIPHDGFLRTPPLIRTOKELSBKIQLEALGDIETAIKLVKTELQSPHEPLDQ 347  
Db 301 acnefytriphdfgrtppllrtqkelskqiqllealgieaikaiklvktelqsphepldq 360  
QY 348 HYRNHLCALRPLDHSYEFKVISQYLOSTHAPTHSDYTMTLDLDFEVKDEKEAFREDL 407  
|||||  
Db 361 hyrnlhcalrpldhsyefkvisyqlsthapthsdymtlldlfevekdekeafred 420  
QY 408 HNRMLLWHGSRMSNMVGLSHGLRIAPPEAPITGYMFGKGIYFADMSKSNYCFASRLK 467  
|||||  
Db 421 hnrmllwhgsrmsnmvlgshglriappeapitgymfgkgyfadmsksaaycfasrlk 480  
QY 468 NTGLLLLSEVALGQCNELEAPRAEGLLGKHSKGLGKMAPSSAHFVTL 518  
|||||  
Db 481 htgllllsevalgqcnelleapraeagllgkhsckglgkmapssahfvtl 531

RESULT 9  
ID AAU29022 standard; Protein; 522 AA.  
XX AAU29022;  
XX  
XX 18-DEC-2001 (first entry)  
DT Mouse PARP-2 protein.  
XX  
DE Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
XX cytostatic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;  
XX immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
XX  
OS Mus musculus.  
XX  
XX WO200164955-A1.  
XX  
XX 07-SEP-2001.  
PD

XX 01-MAR-2001; 2001WO-US06572.  
XX 02-MAR-2000; 2000US-0517467.  
XX (ISIS-) ISIS PHARM INC.  
PI Popoff I, Cowsett LM;  
XX WPI; 2001-602570/68.  
XX N-PSDB; RAS45597.  
XX Antisense compound useful for treating hyperproliferative,  
XX neurological, inflammatory and autoimmune disorders and diabetes  
XX inhibits human PARP -  
XX Example 13; Page 109-111; 168pp; English.  
XX The invention relates to antisense oligonucleotides targeted to human  
XX PARP nucleic acid and inhibiting expression of human PARP. PARP  
XX (Poly (ADP-ribose) polymerase plays an important role in chromatin  
XX decondensation, DNA replication, DNA repair, gene expression, malignant  
XX transformation, cellular differentiation and apoptosis. The antisense  
XX oligonucleotide inhibitors are useful for inhibiting the expression of  
XX PARP in human cells or tissues. They are also useful for treating a  
XX human with a disease associated with PARP especially hyperproliferative  
XX disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
XX neurological (e.g. parkinsonism, meningitis-associated intracranial  
XX complications and ischaemia), inflammatory and autoimmune disorders (e.g.  
XX arthritis) and diabetes. The present sequence is a PARP protein,  
XX the cDNA encoding which was used to design the antisense  
XX oligonucleotides.  
SQ Sequence 522 AA;  
Query Match 82.3%; Score 2468; DB 22; Length 522;  
Best Local Similarity 87.6%; Pred. No. 4.7e-225;  
Matches 465; Conservative 35; Mismatches 21; Indels 10; Gaps 4;  
Qy 40 RRCQRESKKPVAGGKANKORTEDKQDESVALKLLKGGKAPVDECTAKVGAHVYCEGN 99  
Db 2 rtdqr---kgpmaggk-dadrtkdndr-svktlllkgkpvdpcecaaklgahvycgd 55  
Qy 100 DVIDVMLNOTNLQFNNNKYIQLLEDADDQNRNFSVMMRWGRVGMGSHLVACSGNLNKA 159  
Db 56 dvydvmlnqtulqfnmkyiqlledaddqnrnfsvmmrwgrvgtqgshlvtsqgdlnka 115  
Qy 160 KEIFQKKFLDKTKNNWEDREKFEKVPKYDMLQMDYATNTODEEETKKEESLKSPKPES 219  
Db 116 keifqkkfldtknnwedrenfekvpkydmlqmdyaastqdesktkeet----lkpes 171  
Qy 220 OLDLRVQLIKLIGNVQAEEMEMKNTKAPLGKLTVAQIKVAGYOSLKKEIEDCIRAG 279  
Db 172 qldlrvgellkicnvqmeemlemkydtkraplgkltvaqikagyslkkeiedcirag 231  
Qy 280 OHGRALMEACNEFYTRIPDHGLRTPPLIRITQKELSEKIQLEALGDIEIAIKLVKTELQ 339  
Db 232 qhgralveacnefytriphdfglrsippvirkelsdkvkllealgdieialkivksrq 291  
Qy 340 SPEHPLDQHYRNHLCALRPLDHSYEFKVISQYLQSTHAPTHSDYTMTLTLLDFEVEKDG 399  
Db 292 glehpldqhyrnlhcalrpldhsnefkvisqylqsthapthkdytmlltdvfevekege 351  
Qy 400 KEAPREDLHNRMLLWHGRSMNWVILSHGLRIAPPAPITGYWFGKGIYFADMSKSSAN 459  
Db 352 keafredlpmrmlhwhgrslwnvwvlgshglrvappeapitgywfgkgyifadmskssan 411  
Qy 460 YCFASRLKNTGLLLSEVALGQCNLEELANPKAEGLLQGHSTKGLGKMAPSSAHFVTLN 519  
Db 412 ycfasrlkntglllsevalgqcnleelanpkagllrgkstkmgkmapspahftln 471  
Qy 520 GSTVPLGPASDTGILNPDGYTLNTNEXIVVNPQNVRMYLLKVFQNFLLQLW 570

Db 472 gstvplgpasdtgilnpegtylnynefvspnqvmryllkqifnqlw 522  
RESULT 10  
AAB47032  
ID AAB47032 standard; Protein; 1063 AA.  
XX  
AC AAB47032;  
XX  
DT 29-MAR-2001 (first entry)  
XX  
DE Fusion protein PARP1A/PARP2B.  
XX  
KW Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;  
KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;  
KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;  
KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;  
KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;  
KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;  
KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;  
KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;  
KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;  
KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;  
KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;  
KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;  
KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;  
KW graft versus host disease; allograft rejection; cystic fibrosis;  
KW chronic glomerulonephritis; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;  
KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;  
KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;  
KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;  
KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;  
KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;  
KW expressed sequence tag; EST; RACE; PCR; amplify; primer;  
KW polymerase chain reaction.  
OS Synthetic.  
XX  
XX WO200077179-A2.  
XX  
XX 21-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16629.  
XX  
XX 16-JUN-1999; 99US-0139543.  
XX  
XX (ICOS-) ICOS CORP.  
XX  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
XX  
XX WPI; 2001-025335/03.  
XX N-PSDB; AAC85341.  
XX  
XX New human poly(ADP-ribose) polymerase for treating inflammatory,  
XX neurological, cardiovascular, or neoplastic tissue growth disorders,  
XX such as, arthritis, encephalitis, myocardial ischemia, and leukocyte  
XX metastasis -  
XX  
XX Example 3; Page 117-120; 129pp; English.  
XX  
XX This sequence represents the fusion protein PARP1A/PARP2B. This  
XX protein contains amino acids 1-662 of hPARP1 fused upstream of  
XX amino acids 230-583 of hPARP2. The cDNA encoding this sequence was  
XX amplified using the primer sequences given in AAC85321-40 and  
XX AAC85342-51. The fusion protein coding sequence was cloned in a  
XX baculovirus expression vector for the expression of the fusion protein.  
XX The protein of the invention, hPARP2, causes the covalent addition of  
XX polymers of ADP-ribose to protein targets. hPARP2 activity is induced  
XX in many instances of oxidative stress or during inflammation where  
XX there is direct damage to the DNA. hPARP2 may be used to identify  
XX antagonists which may be used to treat a human having a disorder



PR 29-SEP-2000; 2000US-0236369.  
PR 29-SEP-2000; 2000US-0236370.  
PR 02-OCT-2000; 2000US-0236802.  
PR 02-OCT-2000; 2000US-0237037.  
PR 02-OCT-2000; 2000US-0237038.  
PR 02-OCT-2000; 2000US-0237039.  
PR 02-OCT-2000; 2000US-0237040.  
PR 13-OCT-2000; 2000US-0239935.  
PR 13-OCT-2000; 2000US-0239937.  
PR 20-OCT-2000; 2000US-0240960.  
PR 20-OCT-2000; 2000US-0241221.  
PR 20-OCT-2000; 2000US-0241785.  
PR 20-OCT-2000; 2000US-0241786.  
PR 20-OCT-2000; 2000US-0241787.  
PR 20-OCT-2000; 2000US-0241808.  
PR 20-OCT-2000; 2000US-0241809.  
PR 01-NOV-2000; 2000US-0241826.  
PR 01-NOV-2000; 2000US-0244617.  
PR 08-NOV-2000; 2000US-0246474.  
PR 08-NOV-2000; 2000US-0246475.  
PR 08-NOV-2000; 2000US-0246476.  
PR 08-NOV-2000; 2000US-0246477.  
PR 08-NOV-2000; 2000US-0246478.  
PR 08-NOV-2000; 2000US-0246523.  
PR 08-NOV-2000; 2000US-0246524.  
PR 08-NOV-2000; 2000US-0246525.  
PR 08-NOV-2000; 2000US-0246526.  
PR 08-NOV-2000; 2000US-0246527.  
PR 08-NOV-2000; 2000US-0246528.  
PR 08-NOV-2000; 2000US-0246532.  
PR 08-NOV-2000; 2000US-0246609.  
PR 08-NOV-2000; 2000US-0246610.  
PR 08-NOV-2000; 2000US-0246611.  
PR 08-NOV-2000; 2000US-0246613.  
PR 17-NOV-2000; 2000US-0249207.  
PR 17-NOV-2000; 2000US-0249208.  
PR 17-NOV-2000; 2000US-0249209.  
PR 17-NOV-2000; 2000US-0249210.  
PR 17-NOV-2000; 2000US-0249211.  
PR 17-NOV-2000; 2000US-0249212.  
PR 17-NOV-2000; 2000US-0249213.  
PR 17-NOV-2000; 2000US-0249214.  
PR 17-NOV-2000; 2000US-0249215.  
PR 17-NOV-2000; 2000US-0249216.  
PR 17-NOV-2000; 2000US-0249217.  
PR 17-NOV-2000; 2000US-0249218.  
PR 17-NOV-2000; 2000US-0249244.  
PR 17-NOV-2000; 2000US-0249245.  
PR 17-NOV-2000; 2000US-0249264.  
PR 17-NOV-2000; 2000US-0249265.  
PR 17-NOV-2000; 2000US-0249297.  
PR 17-NOV-2000; 2000US-0249299.  
PR 17-NOV-2000; 2000US-0249300.  
PR 01-DEC-2000; 2000US-0250160.  
PR 01-DEC-2000; 2000US-0250391.  
PR 05-DEC-2000; 2000US-0251030.  
PR 05-DEC-2000; 2000US-0251988.  
PR 05-DEC-2000; 2000US-0256719.  
PR 06-DEC-2000; 2000US-0251479.  
PR 08-DEC-2000; 2000US-0251856.  
PR 08-DEC-2000; 2000US-0251868.  
PR 08-DEC-2000; 2000US-0251869.  
PR 08-DEC-2000; 2000US-0251989.  
PR 08-DEC-2000; 2000US-0251990.  
PR 11-DEC-2000; 2000US-0254097.  
PR 05-JAN-2001; 2001US-0259678.  
XX (HUMA-) HUMAN GENOME SCI INC.  
PA Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465558/50.  
DR N-PSDB; AAS34886.

XX Novel polypeptides and polynucleotides useful as diagnostic reagents to  
PT diagnose diseases or disorders associated with aberrant expression or  
PT activity of polypeptides, and for treating cancers, rheumatoid  
PT arthritis  
XX  
PS Claim 11; SEQ ID No 414; 687pp; English.  
XX  
CC The present invention relates to the isolation of novel human neoplastic  
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful in the diagnosis, treatment, prevention and/or  
CC prognosis of disorders involving neoplastic disease such as  
CC hyperproliferative disorders (e.g. leukemia, bone cancer, bladder  
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar  
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may  
CC also be useful for treating other disorders such as neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC and renal disorders. The polynucleotide sequences of the invention are  
CC also useful in gene therapy. AAU21568-AAU21851 represent the novel human  
CC neoplastic disease associated polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 379 AA;

Query Match 65.6%; Score 1966; DB 22; Length 379;  
Best Local Similarity 98.9%; Pred. No. 1.3e-177;  
Matches 375; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 192 QMDYATNTQDEETKESLSKPLAPESQDLRVQELIKLICNVQAMEMEMKNTKK 251  
Db : |||||  
1 exdyatntqdeetkkeslskplkpesqdlrvqeliklicnvqameemmemkntkk 60  
QY 252 APLGKLTVAQIKAGYQSLKKIEDCIRAGOHGRALMEACNEFYTRIPDFGLRTPPLIRTO 311  
Db |||||  
61 aplgkltvaqikagyslkkiedciragohgralmeacnefytriphdfglrtpplirtq 120  
QY 312 KELSEKIOLLEALGDIEIAIKLVKTELQSPHEPLDQHYRNLCALRPLDHSYEFKVISQ 371  
Db |||||  
121 kelxekiqllealgdieiaiklvktelqsphepldqhyrnlhcalrpldhesyefkvisq 180  
QY 372 YLOSTHAPTHSDYTWLIDLFEVEKDGKEAPREDLHNPMLLWHGSRMSNWGILSHGLR 431  
Db |||||  
181 ylosthapthsdytntlidlfevekdgekeafredlhnrmllwhgsrmsnwvgilshglr 240  
QY 432 IAPPEAPITGYMFGKIYFADMSKSNYCFASRLKNTGLLLSEVALGQCNELEFANPK 491  
Db |||||  
241 iappeapitgymfgkiyfadmssksnycfaskrknktglllsevalgqcneleapnk 300  
QY 492 AEGLLQGHSTKGLKMAPSSAHFVTLNGSTVPLGPASDTGLNPDGYTLNVEIVNYP 551  
Db |||||  
301 aegllqghstkglgkmapssahfvtlngstvtplgpasdtglnpdgtylnveyivnyp 360  
QY 552 NOVNRMYLLKQFNFQLQW 570  
Db |||||  
361 novnrmyllkqfnflqlw 379

RESULT 12  
AAB47030

ID AAB47030 standard; Protein; 360 AA.

XX AAB47030;

AC AAB47030;

XX 29-MAR-2001 (first entry)

DT N-terminal fragment of hPARP2.

XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;  
KW

KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;  
KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;  
KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;  
KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;  
KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;  
KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;  
KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;  
KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;  
KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;  
KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;  
KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;  
KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;  
KW graft versus host disease; allograft rejection; cystic fibrosis;  
KW chronic glomerulonephritis; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;  
KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;  
KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;  
KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;  
KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;  
KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;  
KW expressed sequence tag; EST; RACE; PCR; amplify; primer;  
KW polymerase chain reaction.  
XX  
OS Synthetic.  
XX  
XX WO200077179-A2.  
XX  
XX 21-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16629.  
XX  
XX 16-JUN-1999; 99US-0139543.  
XX  
XX (ICOS-) ICOS CORP.  
XX  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
PI  
XX  
XX WPI; 2001-025335/03.  
XX  
XX N-PSDB; AAC85314.  
XX  
XX New human poly(ADP-ribose) polymerase for treating inflammatory,  
PT neurological, cardiovascular, or neoplastic tissue growth disorders,  
PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte  
PT metastasis -  
XX  
PS Example 2; Page 100-01; 129pp; English.  
XX  
CC This sequence is encoded by the 5' fragment of the human parp2  
CC (poly(ADP-ribose) polymerase) cDNA. The cDNA sequence was amplified  
CC using the primers given in AAC85307-13. DNA derived from spleen and  
CC testis cDNA libraries were used as templates. The primers given  
CC in AAC85315-17 were used to confirm the sequence of the 5'-hPARP2  
CC fragment. The amplified fragments, P2-1 and P2-9 (AAC85318-19) were  
CC cloned to help determine that hPARP2 has an open reading frame of  
CC 1080 nucleotides. The protein of the invention, hPARP2, causes  
CC the covalent addition of polymers of ADP-ribose to protein targets.  
CC hPARP2 activity is induced in many instances of oxidative stress or  
CC during inflammation where there is direct damage to the DNA.  
CC hPARP2 may be used to identify antagonists which may be used to  
CC treat a human having a disorder mediated by PARP2 activity, such as,  
CC inflammatory, neurological, cardiovascular, or neoplastic tissue  
CC growth disorders. hPARP2 and antibodies to it, can also be used  
CC to diagnose these conditions.  
XX  
SQ Sequence 360 AA;

Query Match 59.7%; Score 1790.5; DB 22; Length 360;  
Best local Similarity 96.4%; Pred. No. 5.4e-161;  
Matches 347; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
1 MAARRRRSTGGGRARALNESKRVNNGNTAPEDSSPAKTRRCQRESKMPVAGGKANKD 60  
|||||

Db 1 maarrtrstgggraralneskrvnngntapedsspaktrrcqrqeskkmpvaggkankd 60  
QY 61 RTEDKQD-----ESVKALLLKGKAPVDPDECTAKVGAHVYCEGNDVYDMLN 107  
Db 61 rtedkqdgmpgrswaskrvsesvkalllkgkapvdpdectakvgkahnvcegnvdydmin 120  
QY 108 QTNLFNNNNYYLIQLLEDDAQRNFVNMWRGVRGKMGQHSLVACSGNLNKAKEIFQKFF 167  
Db 121 qtnlfnnnnkyyliqlleddaqrnfsvnmwrgrvrgkmgqhslvacsgnlnkakeifqkff 180  
QY 168 LDKTKNNWEDREKFEKVPKGYDMLQMDYATNTQDEETKKEESLKSPLKPESOLDLRVOE 227  
Db 181 ldktknnwredrekfekvpkgkymldmqdyatntqdeetkkeslksplkpesqldlrve 240  
QY 228 LKLIKCNVQAEEMMEMKYNFKKAPLGKLTVAQIKAGYQSLKKIEDCIKAGOHGRALME 287  
Db 241 lklikcnvqaeemmemkynfkckaplgkltvaqikagyslkkiedcirkagohgralne 300  
QY 288 ACNEFYTRIPHPDGLTTPPLIRTQKELSEKIQLEALGDIETAIKLVKTELSPHEPLDQ 347  
Db 301 acnefytriphdfgirtpplirtqkelsekiqllealgieaiklvktelqsphepldq 360

RESULT 13  
AAB47031  
ID AAB47031 standard; Protein; 287 AA.  
AC AAB47031;  
XX  
XX 29-MAR-2001 (first entry)  
XX  
XX C-terminal fragment of hPARP2.  
XX  
XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;  
KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;  
KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;  
KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;  
KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;  
KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;  
KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;  
KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;  
KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;  
KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;  
KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;  
KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;  
KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;  
KW graft versus host disease; allograft rejection; cystic fibrosis;  
KW chronic glomerulonephritis; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;  
KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;  
KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;  
KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;  
KW hypovolemic shock; Type 1 diabetes mellitus; hypersensitivity;  
KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity;  
KW expressed sequence tag; EST; RACE; PCR; amplify; primer;  
KW polymerase chain reaction.  
XX  
OS Synthetic.  
XX  
XX WO200077179-A2.  
XX  
XX 21-DEC-2000.  
XX  
XX 16-JUN-2000; 2000WO-US16629.  
XX  
XX 16-JUN-1999; 99US-0139543.  
XX  
XX (ICOS-) ICOS CORP.  
XX  
XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
PI  
XX  
XX WPI; 2001-025335/03.  
XX  
XX N-PSDB; AAC85320.

xx New human poly(ADP-ribose) polymerase for treating inflammatory,  
PT neurological, cardiovascular, or neoplastic tissue growth disorders,  
PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte  
PT metastasis -  
xx

Example 2; Page 101-02; 129pp; English.

This sequence is encoded by the 3' fragment of the human parp2  
(poly(ADP-ribose) polymerase) cDNA. The cDNA sequence was amplified  
using the primers given in AAC85315-17. DNA derived from a testis  
cDNA library was used as a template. The protein of the invention,  
hPARP2, causes the covalent addition of polymers of ADP-ribose to  
protein targets. hPARP2 activity is induced in many instances of  
oxidative stress or during inflammation where there is direct  
damage to the DNA. hPARP2 may be used to identify antagonists  
which may be used to treat a human having a disorder mediated by  
PARP2 activity, such as, inflammatory, neurological, cardiovascular,  
or neoplastic tissue growth disorders. hPARP2 and antibodies to it,  
can also be used to diagnose these conditions.

xx Sequence 287 AA;

Query Match 50.6%; Score 1518; DB 22; Length 287;  
Best Local Similarity 100.0%; Pred. No. 2.8e-135;  
Matches 287; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 284 ALMEACNEFYTRIPHDGRLPPLIRTOKEISEKIQLEALGDIEIAIKLVKTELQSPH 343  
Db 1 almeacnefytriphdfgrltpplirtqeksekiqllealgdieialkvkltelqsp 60  
Qy 344 PLDQHYRNLCALRPLDHESEFEKVISQYLOSTHAPTHSDYTMTLDDLFEVKDGEKEAF 403  
Db 61 pldqhyrnlhcalrpldhesyefkvisyldsthapthsdymtldldifevkegekeaf 120  
Qy 404 REDLHNRMLLWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKGIYFADMSKSNANYCFA 463  
Db 121 redlhnrlmlwhgsrmsnwgilshglriappeapitgymfgkgyifadmsksanycfa 180  
Qy 464 SRLKNTGILLSEVALGQCNELLENPKAEGLLQGHKSTKGLGKMAPSSAHFVILNGSTV 523  
Db 181 srlkntgilllsevalgqcnelleanpkagllqghkstkglgkmapssahfvilngstv 240  
Qy 524 PLGPASDTGILNPDGTYTLYNNEYIYVNPQVRYMRYLLKXQENFLQLW 570  
Db 241 plgpasdtgillnpgdytlynneyivnqvmryllkvqfnfqlw 287

RESULT 14

AAU20129

ID AAU20129 standard; Protein; 294 AA.

xx

xx

xx

xx

xx

xx Human DNA repair and processing polypeptide #14.  
xx DNA processing; human; mouse; rabbit; goat; horse; cat; gene therapy;  
xx chicken; sheep; immunosuppressive; antiarthritic; vasotropic;  
xx antirheumatic; antiproliferative; cytostatic; cardiant; neuroprotective;  
xx cerebrotective; nootropic; antibacterial; virucide; fungicide; cancer;  
xx ophthalmological; vulnery; gene therapy; autoimmune disease; neoplasm;  
xx hyperproliferative disorder; breast; liver; cardiovascular disorder;  
xx cerebrovascular disorder; nervous system disorder; bacterial infection;  
xx fungal infection; viral infection; ocular disorder; endocrine disorder;  
xx gastrointestinal disorder; renal disorder; respiratory disorder; dog;  
xx wound healing; skin aging; organ transplantation; tissue regeneration;  
xx anti-infertility; DNA repair protein.

OS Homo sapiens.

xx

PN WO20015204-A1.  
xx  
xx  
PD 02-AUG-2001.  
xx  
xx  
PF 17-JAN-2001; 2001WO-US01336.  
xx  
xx  
PR 31-JAN-2000; 2000US-0179065.  
PR 04-FEB-2000; 2000US-0180628.  
PR 24-FEB-2000; 2000US-0184664.  
PR 02-MAR-2000; 2000US-0186350.  
PR 16-MAR-2000; 2000US-0189874.  
PR 17-MAR-2000; 2000US-0190076.  
PR 18-APR-2000; 2000US-0198123.  
PR 19-MAY-2000; 2000US-0205515.  
PR 07-JUN-2000; 2000US-0209467.  
PR 28-JUN-2000; 2000US-0214886.  
PR 30-JUN-2000; 2000US-0215135.  
PR 07-JUL-2000; 2000US-0216647.  
PR 07-JUL-2000; 2000US-0216880.  
PR 11-JUL-2000; 2000US-0217487.  
PR 11-JUL-2000; 2000US-0217496.  
PR 14-JUL-2000; 2000US-0218290.  
PR 26-JUL-2000; 2000US-0220963.  
PR 26-JUL-2000; 2000US-0220964.  
PR 14-AUG-2000; 2000US-0224518.  
PR 14-AUG-2000; 2000US-0224519.  
PR 14-AUG-2000; 2000US-0225213.  
PR 14-AUG-2000; 2000US-0225214.  
PR 14-AUG-2000; 2000US-0225266.  
PR 14-AUG-2000; 2000US-0225267.  
PR 14-AUG-2000; 2000US-0225268.  
PR 14-AUG-2000; 2000US-0225270.  
PR 14-AUG-2000; 2000US-0225447.  
PR 14-AUG-2000; 2000US-0225757.  
PR 14-AUG-2000; 2000US-0225758.  
PR 14-AUG-2000; 2000US-0225759.  
PR 18-AUG-2000; 2000US-0226279.  
PR 22-AUG-2000; 2000US-0226681.  
PR 22-AUG-2000; 2000US-0226868.  
PR 22-AUG-2000; 2000US-0227182.  
PR 23-AUG-2000; 2000US-0227009.  
PR 30-AUG-2000; 2000US-0228924.  
PR 01-SEP-2000; 2000US-0229287.  
PR 01-SEP-2000; 2000US-0229343.  
PR 01-SEP-2000; 2000US-0229344.  
PR 01-SEP-2000; 2000US-0229345.  
PR 05-SEP-2000; 2000US-0229509.  
PR 05-SEP-2000; 2000US-0229513.  
PR 06-SEP-2000; 2000US-0230437.  
PR 06-SEP-2000; 2000US-0230438.  
PR 08-SEP-2000; 2000US-0231242.  
PR 08-SEP-2000; 2000US-0231243.  
PR 08-SEP-2000; 2000US-0231244.  
PR 08-SEP-2000; 2000US-0231413.  
PR 08-SEP-2000; 2000US-0231414.  
PR 08-SEP-2000; 2000US-0232080.  
PR 08-SEP-2000; 2000US-0232081.  
PR 12-SEP-2000; 2000US-0231968.  
PR 14-SEP-2000; 2000US-0232397.  
PR 14-SEP-2000; 2000US-0232398.  
PR 14-SEP-2000; 2000US-0232399.  
PR 14-SEP-2000; 2000US-0232400.  
PR 14-SEP-2000; 2000US-0232401.  
PR 14-SEP-2000; 2000US-0233063.  
PR 14-SEP-2000; 2000US-0233064.  
PR 14-SEP-2000; 2000US-0233065.  
PR 21-SEP-2000; 2000US-0234223.  
PR 21-SEP-2000; 2000US-0234274.  
PR 25-SEP-2000; 2000US-0234997.  
PR 25-SEP-2000; 2000US-0234998.  
PR 26-SEP-2000; 2000US-0235484.  
PR 27-SEP-2000; 2000US-0235834.  
PR 27-SEP-2000; 2000US-0235836.

PR 29-SEP-2000; 2000US-0236327.  
 PR 29-SEP-2000; 2000US-0236367.  
 PR 29-SEP-2000; 2000US-0236368.  
 PR 29-SEP-2000; 2000US-0236369.  
 PR 29-SEP-2000; 2000US-0236370.  
 PR 02-OCT-2000; 2000US-0236802.  
 PR 02-OCT-2000; 2000US-0237037.  
 PR 02-OCT-2000; 2000US-0237038.  
 PR 02-OCT-2000; 2000US-0237039.  
 PR 02-OCT-2000; 2000US-0237040.  
 PR 13-OCT-2000; 2000US-0239935.  
 PR 13-OCT-2000; 2000US-0239937.  
 PR 20-OCT-2000; 2000US-0240960.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241221.  
 PR 20-OCT-2000; 2000US-0241785.  
 PR 20-OCT-2000; 2000US-0241786.  
 PR 20-OCT-2000; 2000US-0241787.  
 PR 20-OCT-2000; 2000US-0241808.  
 PR 20-OCT-2000; 2000US-0241809.  
 PR 20-OCT-2000; 2000US-0241826.  
 PR 01-NOV-2000; 2000US-0244617.  
 PR 08-NOV-2000; 2000US-0246474.  
 PR 08-NOV-2000; 2000US-0246475.  
 PR 08-NOV-2000; 2000US-0246476.  
 PR 08-NOV-2000; 2000US-0246477.  
 PR 08-NOV-2000; 2000US-0246478.  
 PR 08-NOV-2000; 2000US-0246523.  
 PR 08-NOV-2000; 2000US-0246524.  
 PR 08-NOV-2000; 2000US-0246525.  
 PR 08-NOV-2000; 2000US-0246526.  
 PR 08-NOV-2000; 2000US-0246527.  
 PR 08-NOV-2000; 2000US-0246528.  
 PR 08-NOV-2000; 2000US-0246532.  
 PR 08-NOV-2000; 2000US-0246539.  
 PR 08-NOV-2000; 2000US-0246609.  
 PR 08-NOV-2000; 2000US-0246610.  
 PR 08-NOV-2000; 2000US-0246611.  
 PR 08-NOV-2000; 2000US-0246613.  
 PR 17-NOV-2000; 2000US-0249207.  
 PR 17-NOV-2000; 2000US-0249208.  
 PR 17-NOV-2000; 2000US-0249209.  
 PR 17-NOV-2000; 2000US-0249210.  
 PR 17-NOV-2000; 2000US-0249211.  
 PR 17-NOV-2000; 2000US-0249212.  
 PR 17-NOV-2000; 2000US-0249213.  
 PR 17-NOV-2000; 2000US-0249214.  
 PR 17-NOV-2000; 2000US-0249215.  
 PR 17-NOV-2000; 2000US-0249216.  
 PR 17-NOV-2000; 2000US-0249217.  
 PR 17-NOV-2000; 2000US-0249218.  
 PR 17-NOV-2000; 2000US-0249244.  
 PR 17-NOV-2000; 2000US-0249245.  
 PR 17-NOV-2000; 2000US-0249264.  
 PR 17-NOV-2000; 2000US-0249265.  
 PR 17-NOV-2000; 2000US-0249297.  
 PR 17-NOV-2000; 2000US-0249299.  
 PR 17-NOV-2000; 2000US-0249300.  
 PR 01-DEC-2000; 2000US-0250160.  
 PR 01-DEC-2000; 2000US-0250391.  
 PR 05-DEC-2000; 2000US-0251030.  
 PR 05-DEC-2000; 2000US-0251988.  
 PR 05-DEC-2000; 2000US-0256719.  
 PR 06-DEC-2000; 2000US-0251479.  
 PR 08-DEC-2000; 2000US-0251856.  
 PR 08-DEC-2000; 2000US-0251868.  
 PR 08-DEC-2000; 2000US-0251869.  
 PR 08-DEC-2000; 2000US-0251989.  
 PR 08-DEC-2000; 2000US-0251990.  
 PR 11-DEC-2000; 2000US-0254097.  
 PR 05-JAN-2001; 2001US-0259678.  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 FA Rosen CA, Barash SC, Ruben SM;  
 XX  
 PI

XX WPI; 2001-496846/54.  
 DR N-PSDB; AAS32196.  
 XX  
 PT Nucleic acids encoding human polypeptides, useful for preventing,  
 PT diagnosing and/or treating e.g. cancers, Parkinson's disease and  
 PT diabetic retinopathy -  
 XX  
 PS Claim 11; SEQ ID No 56; 460pp; English.  
 XX  
 CC Sequences AAU20116-AAU20147 represent the DNA repair and processing  
 CC polypeptides of the invention. DNA repair and processing polypeptides and  
 CC their associated polynucleotides are useful in the diagnosis, treatment  
 CC and prevention of various types of disorders in e.g. humans, mice,  
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. A pathological  
 CC condition can be determined by detecting the presence or absence of a  
 CC mutation in a DNA repair and processing polynucleotide. The treatable  
 CC disorders include autoimmune diseases such as rheumatoid arthritis,  
 CC hyperproliferative disorders such as neoplasms of the breast or liver,  
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular  
 CC disorders such as cerebral ischaemia, nervous system disorders such as  
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,  
 CC ocular disorders such as corneal infection, endocrine disorders such as  
 CC premature labour and infertility, gastrointestinal disorders such as  
 CC Crohn's disease, renal disorders such as glomerulonephritis and  
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can  
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,  
 CC to maintain organs before transplantation, to regenerate tissues and in  
 CC chemotaxis.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic format directly from WIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 Query Match 48.68; Score 1456; DB 22; Length 294;  
 Best Local Similarity 97.9%; Pred. No. 2.2e-129;  
 Matches 280; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 Qy 192 QMDYATNTQDEETKESLSPKSPESQLDLRVQELIKLCNVQAEEMMEMKYNKK 251  
 Db : |||||  
 1 exdyatntqdeetkkeslspkspesqldrvqelikkicnvqaeemmemkynk 60  
 Qy 252 APLGKLTVAQIKAGYQSLKKIEDCIRAGOHGRALMEACNEFYTRIPHPGLTPPLIRFQ 311  
 Db |||||  
 61 aplgkltvaqikagyslkiedciraqghgxalmeacnefytriphdfglrtplrtq 120  
 Qy 312 KELSEKIQLEALGDIEIAIKLVKTELOSPHPDLQHVNLHLCALRPLDHSEYEFKVISQ 371  
 Db |||||  
 121 kelxekiqllealgdieiaiklvktselqsphepdlqhvnlhcalrpldhse yefkvisq 180  
 Qy 372 YLQSTHAPTHSDYTMTLDDLFEVEKDGKEAFREDLHNRMLIWHGSRMSNMYGILSHGLR 431  
 Db |||||  
 181 ylsthaphthsdymtldlfevekdgeafredlhmrlmlwhgsrmsnmygilshglr 240  
 Qy 432 IAPPEAPITGYMFGKIYFADMSKSNANYCFASRLKNTGILLSEV 477  
 Db |||||  
 241 iappeaxitgymxgkiyfadmssksanycfasrikntgillllse 286  
 RESULT 15  
 AAU21810  
 ID AAU21810 standard; Protein: 294 AA.  
 XX  
 XX AAU21810;  
 XX  
 DT 06-DEC-2001 (first entry)  
 XX  
 DE Novel human neoplastic disease associated polypeptide #243.  
 XX  
 KW Human; neoplastic disease associated polypeptide; cancer;  
 KW hyperproliferative disorder; neural disorder; immune system disorder;  
 KW muscular disorder; reproductive disorder; gastrointestinal disorder;  
 KW pulmonary disorder; cardiovascular disorder; renal disorder;  
 KW neuroprotective; cytostatic; anti inflammatory; vasotropic.





PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Rosen CA, Barash SC, Ruben SM;  
XX  
XX WPI; 2001-465558/50.  
DR N-PSDB; AAS35009.  
XX  
PT Novel polypeptides and polynucleotides useful as diagnostic reagents to  
PT diagnose diseases or disorders associated with aberrant expression or  
PT activity of polypeptides, and for treating cancers, rheumatoid  
PT arthritis  
XX  
XX  
PS Claim 11: SEQ ID No 537; 687pp: English.  
XX  
XX The present invention relates to the isolation of novel human neoplastic  
CC disease associated polypeptides, and cDNA (AAS34767-AAS35050) and DNA  
CC sequences encoding for these polypeptides. The sequences of the  
CC invention are useful in the diagnosis, treatment, prevention and/or  
CC prognosis of disorders involving neoplastic disease such as  
CC hyperproliferative disorders (e.g. leukaemia, bone cancer, bladder  
CC cancer, brain stem glioma, adult liver cancer, childhood cerebellar  
CC astrocytoma, or Hodgkin's lymphoma). The sequences of the invention may  
CC also be useful for treating other disorders such as neural disorders,  
CC immune system disorders, muscular disorders, reproductive disorders,  
CC gastrointestinal disorders, pulmonary disorders, cardiovascular disorders  
CC and renal disorders. The polynucleotide sequences of the invention are  
CC also useful in gene therapy. AAU21568-AAU21851 represent the novel human  
CC neoplastic disease associated polypeptides of the invention.  
CC Note: The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic format directly from WIPO  
CC at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 294 AA;

Query Match 48.6%; Score 1456; DB 22; Length 294;  
Best Local Similarity 97.9%; Pred. NO. 2.2e-129;  
Matches 280; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 192 QNDYATNTQDEETKKEESLKSPLKPSQDLRVQELIKLCNVQAMEMMEMKYNTKK 251  
Db : |||||  
1 exdyatntqdeetkkeslksplkpsqdlrvqeliklcnvqameemmemkyntkk 60

OY 252 APLGLTVAQIKAGYQSLKKIEDCIRAGQGRALMEACNEFYTRIPHDGRLTPPLIRTO 311  
Db |||||  
61 aplglstvaikagyslkkiedciragqghxalmeacnefytriphdglrtplirtq 120

OY 312 KELSEKIQLEALGDIEITAIVKTELQSPHEPLDQHYRNHLCALRPLDHSYEFKVISQ 371  
Db |||||  
121 kelxekiqllealgdieitaikvktelqsphepldqghyrnlhcalrpldhsyefkvisq 180

OY 372 YLQSTHAPTHSDYTWTLTDLFEVEKDGKEAFREDLHNRMLLWHGSRMSNWVGIILSHGLR 431  
Db |||||  
181 ylqsthapthsdytwtltdlfevekdgekeafredlhnrmllwhgsrmsnwvgiilshglr 240

OY 432 IAPPEAPITGYNFGYFADMSKSNYCFASRLKNTGLLLSEV 477  
Db |||||  
241 iappeaxitgymxgkyfadmssksnycfasrlkntgllllse 286

Search completed: August 29, 2002, 07:57:12  
Job time: 289 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:53:23 ; Search time 46.79 Seconds  
(without alignments)  
297.555 Million cell updates/sec

Title: US-09-701-586B-2

Perfect score: 2998

Sequence: 1 MAARRRSTGGGRARALNES.....PNQVRMYLLKVFNFQLQW 570

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCIRUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1012.5	33.8	1013	4	US-08-860-886-2
2	1007.5	33.6	1014	4	US-09-078-347A-3
3	135	4.5	1327	4	US-09-196-387-2
4	120.5	4.0	469	4	US-09-052-089A-1
5	119	4.0	163	1	US-08-044-618-5
6	116.5	3.9	469	2	US-08-968-751-2
7	115	3.8	906	2	US-08-609-230A-9
8	112.5	3.8	560	1	US-08-336-618-22
9	112.5	3.8	1312	2	US-08-687-080-51
10	111	3.7	3218	1	US-08-764-100-27
11	110.5	3.7	700	1	US-07-720-589-2
12	110.5	3.7	700	2	US-08-785-190-2
13	110.5	3.7	700	4	US-08-235-836C-66
14	110.5	3.7	700	5	PCT-US92-05539-2
15	110.5	3.7	907	3	US-08-990-140-4
16	110.5	3.7	907	4	US-09-546-238-4
17	110.5	3.7	940	4	US-08-810-712-7
18	110.5	3.7	1312	2	US-08-592-126-148
19	110	3.7	683	6	5210183-3
20	109.5	3.7	693	4	US-08-235-836C-68
21	109	3.6	926	3	US-08-755-587-187
22	108.5	3.6	807	4	US-09-081-345-2
23	108	3.6	458	1	US-08-336-618-24
24	107.5	3.6	619	1	US-08-465-746-2
25	107.5	3.6	619	1	US-08-214-164-2
26	107.5	3.6	619	2	US-08-467-852A-3
27	107.5	3.6	619	2	US-08-246-636-2

28	107.5	3.6	619	2	US-08-247-491A-3	Sequence 3, Appli
29	107.5	3.6	619	2	US-08-319-795-2	Sequence 2, Appli
30	107.5	3.6	619	2	US-08-468-985-2	Sequence 2, Appli
31	107.5	3.6	619	3	US-08-312-949-2	Sequence 2, Appli
32	107.5	3.6	648	1	US-08-072-070-2	Sequence 2, Appli
33	107.5	3.6	648	1	US-08-469-434-2	Sequence 2, Appli
34	107.5	3.6	648	1	US-08-214-222-2	Sequence 2, Appli
35	107.5	3.6	648	2	US-08-467-852A-2	Sequence 2, Appli
36	107.5	3.6	648	2	US-08-468-718-2	Sequence 2, Appli
37	107.5	3.6	648	2	US-08-247-491A-2	Sequence 2, Appli
38	107.5	3.6	648	3	US-08-446-201-3	Sequence 3, Appli
39	107.5	3.6	695	1	US-08-127-499A-23	Sequence 23, Appli
40	107.5	3.6	695	1	US-08-482-847-23	Sequence 23, Appli
41	106.5	3.6	441	4	US-08-630-915A-34	Sequence 34, Appli
42	106.5	3.6	700	4	US-08-235-836C-74	Sequence 74, Appli
43	106	3.5	470	4	US-09-052-089A-2	Sequence 2, Appli
44	106	3.5	693	4	US-08-235-836C-72	Sequence 72, Appli
45	106	3.5	2482	1	US-08-328-254-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-08-860-886-2  
; Sequence 2, Application US/08860886  
; Patent No. 6335009  
; GENERAL INFORMATION:  
; APPLICANT: Burkli, Alexander  
; APPLICANT: Zur Hausen, Harald  
; APPLICANT: Jan-Heiner, Kupper  
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE  
; TITLE OF INVENTION: IN GENE THERAPY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds, LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/860,886  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 8484-0028-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-860-886-2

Query Match 33.8%; Score 1012.5; DB 4; Length 1013;

Best Local Similarity 40.5%; Pred. No. 6.6e-89;  
Matches 221; Conservative 111; Mismatches 185; Indels 29; Gaps 11;

Qy 29 APEDSSPAKTRRCORQESKKMPVAGGKANKDRTEKQDESVKALLLKGKAPVDPCTAK 88  
Db 493 APRKSGAALSKKSGQV-----KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 538  
Qy 89 VGKAHYCEGNDVYDVMNLQTNFNKNNKYYLIQLLEDDAQRNFVVMWGRVCKMGQHS 148  
Db 539 EHSAHVLEKGGKVSATLGLVDIVKGTNSYYKQLLEDDKENRYWIFRSRGVGTIGSN 598  
Qy 149 LVACSGNLNKAKEIFQKFLDKTKNNEDREKFEKYPGKYDMLQMDYATNTQDEETKKE 208  
Db 599 KLEOMPSKEDAIEHFMKLYEKTGNAMHSK-NFTKYPKFYPEIDYG---QDEEAVKK- 653  
Qy 209 ESLSKPLKPSQDLRVOELIKLNCVOAMEEMMKYNTKKAPLGLTVAQIKAGYQS 268  
Db 654 --LTVNPGTKSKLPKQVODLIKMFVDSMKKAMVEYIDLQKMPGLKSKRQIQAAYSI 711  
Qy 269 LKKTEDCIRAGOHGRALMEACNEFYTRIPHDGLRTPLIRTQKELSEKIQLEALGDIE 328  
Db 712 LSEVQAVQSSQSDLSNRYTLIPHDGFMKKPPLNNADSVQAKVEMLDNLDDIE 771  
Qy 329 IATLKV-TELQSPHPLDQHYRNLCALRPLDHSYEFKVISQYLOSTHAPTHTSDYMT 387  
Db 772 VAYSLLRGSDSDSKDPIDVNYEKLKTDIKVDRDSEAEIIRKYVKNTHATTNAYDLE 831  
Qy 388 LLDLFEVEKDEKEAFR--EDLHNRMLLWHSRMSNVGILSHGLRIAPPEAPITGYMFG 445  
Db 832 VIDIFKIEREGECORYKPFQKLNHRLLWHSRTNFAGILSQGLRIAPPEAPITGYMFG 891  
Qy 446 KGIYFADMSKSNYCFASRLKNTGLLLSEVALGQCNEELLEAPNKAEGLLQGHSTKGL 505  
Db 892 KGIYFADMSKSNYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPKGKHSVKGL 950  
Qy 506 GKMAPSSAHEVTNGSTVPLGSPASDTGLNPDGYTLNAYNEIVYNPNQVRMYLLKQVFN 565  
Db 951 GKTPDPDSANISLDGVDVPLGTGSSGV--NDTSLLYNEIVYDIAQVNLKYLKLUKFN 1007  
Qy 566 F-LQLW 570  
Db 1008 FKTSLW 1013

RESULT 2  
US-09-078-347A-3  
; Sequence 3, Application US/09078347A  
; Patent No. 6132968  
; GENERAL INFORMATION:  
; APPLICANT: Le, Xiao-Chun  
; APPLICANT: Weinfeld, Michael  
; APPLICANT: Xing, James Z.  
; TITLE OF INVENTION: Methods for Quantitating Low Level  
; TITLE OF INVENTION: Modifications of Nucleotide Sequences  
; NUMBER OF SEQUENCES: 3  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/078,347A  
; FILING DATE: 13-MAY-1998  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Carroll, Peter G.

REGISTRATION NUMBER: 32,837  
REFERENCE/DOCKET NUMBER: UALB-03283  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 705-8410  
TELEFAX: (415) 397-8338  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1014 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-09-078-347A-3

Query Match 33.6%; Score 1007.5; DB 4; Length 1014;  
Best Local Similarity 40.5%; Pred. No. 2e-88;  
Matches 221; Conservative 109; Mismatches 187; Indels 29; Gaps 11;

Qy 29 APEDSSPAKTRRCORQESKKMPVAGGKANKDRTEKQDESVKALLLKGKAPVDPCTAK 88  
Db 494 APRKSGAALSKKSGQV-----KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 539  
Qy 89 VGKAHYCEGNDVYDVMNLQTNFNKNNKYYLIQLLEDDAQRNFVVMWGRVCKMGQHS 148  
Db 540 EHSAHVLEKGGKVSATLGLVDIVKGTNSYYKQLLEDDKENRYWIFRSRGVGTIGSN 599  
Qy 149 LVACSGNLNKAKEIFQKFLDKTKNNEDREKFEKYPGKYDMLQMDYATNTQDEETKKE 208  
Db 600 KLEOMPSKEDAIEHFMKLYEKTGNAMHSK-NFTKYPKFYPEIDYG---QDEEAVKK- 654  
Qy 209 ESLSKPLKPSQDLRVOELIKLNCVOAMEEMMKYNTKKAPLGLTVAQIKAGYQS 268  
Db 655 --LTVNPGTKSKLPKQVODLIKMFVDSMKKAMVEYIDLQKMPGLKSKRQIQAAYSI 712  
Qy 269 LKKTEDCIRAGOHGRALMEACNEFYTRIPHDGLRTPLIRTQKELSEKIQLEALGDIE 328  
Db 713 LSEVQAVQSSQSDLSNRYTLIPHDGFMKKPPLNNADSVQAKVEMLDNLDDIE 772  
Qy 329 IATLKV-TELQSPHPLDQHYRNLCALRPLDHSYEFKVISQYLOSTHAPTHTSDYMT 387  
Db 773 VAYSLLRGSDSDSKDPIDVNYEKLKTDIKVDRDSEAEIIRKYVKNTHATTNAYDLE 832  
Qy 388 LLDLFEVEKDEKEAFR--EDLHNRMLLWHSRMSNVGILSHGLRIAPPEAPITGYMFG 445  
Db 833 VIDIFKIEREGECORYKPFQKLNHRLLWHSRTNFAGILSQGLRIAPPEAPITGYMFG 892  
Qy 446 KGIYFADMSKSNYCFASRLKNTGLLLSEVALGQCNEELLEAPNKAEGLLQGHSTKGL 505  
Db 893 KGIYFADMSKSNYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPKGKHSVKGL 951  
Qy 506 GKMAPSSAHEVTNGSTVPLGSPASDTGLNPDGYTLNAYNEIVYNPNQVRMYLLKQVFN 565  
Db 952 GKTPDPDSANISLDGVDVPLGTGSSGVIDT---SLLYNEIVYDIAQVNLKYLKLUKFN 1008  
Qy 566 F-LQLW 570  
Db 1009 FKTSLW 1014

RESULT 3  
US-09-196-387-2  
; Sequence 2, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/196,387  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
US-09-196-387-2

Query Match 4.5%; Score 135; DB 4; Length 1327;

Best Local Similarity 24.4%; Pred. NO. 0.001;

Matches 75; Conservative 31; Mismatches 104; Indels 98; Gaps 15;

QY 312 KELSEKIQL-LEALGDI-----ETAI-----KLVK-----TELQSPHEPLDOHYRNL 352  
DB 1044 RDIFETEQTLDVADMGMHEELKEIGINAYHRHKLIRKVERLLGGQGGTNP-----YLTFF 1099  
QY 353 HCA-----LRPLDHSYEFKVISQYLSQSTHAPTHSD-----YTMTLDLDFEVEKD 397  
DB 1100 HCVNGTITLLDAPEDKEYQSVVEEMQST-IREHRDGGNAGIFNRYNVIQKVYNNKL 1158  
QY 398 GEKEAFRE-----DLNHRMLLWHGSRMSNWVGILSHGLRIAPPEAPITGYMFGKGIY 449  
DB 1159 RERFCHROKEVSEENHNHNRMLPHGSPFIN--AIHKGf---DERHAYIGMGFGAGIY 1213  
QY 450 FADMSKSNAYCFASRLKNTGL-----LLSEVALQCQNELLEANPKASGL 495  
DB 1214 FAENSSKNQYVYIGT-GGTGCTPHKDRSCYICHRQMLFCRVTLGK----- 1258  
QY 496 LQGHSTKGLKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNXYNIYVNPNOVR 555  
DB 1259 -----SFLQFSTMKMAHAPPGHHSVIG--RPSVNGLAYAEVYIRGEQAY 1301  
QY 556 MRYLLKVQ 563  
DB 1302 PEYLITYQ 1309

RESULT 4

US-09-052-089A-1

; Sequence 1, Application US/09052089A

; Patent No. 6346605

; GENERAL INFORMATION:

; APPLICANT: Lee, Soo Y.

; Choi, Yongwon

; TITLE OF INVENTION: SIGNAL TRANSDUCER FOR THE TNF RECEPTOR SUPER  
FAMILY, AND USES THEREOF

NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: David A. Jackson, Esq.  
STREET: 411 Hackensack Ave, Continental Plaza, 4th  
Floor  
CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/052,089A  
FILING DATE: 31-Mar-1998  
CLASSIFICATION: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-198 CIP 1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 469 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO  
FRAGMENT TYPE: <Unknown>  
ORIGINAL SOURCE:  
ORGANISM: Homo sapiens  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-052-089A-1

Query Match 4.0%; Score 120.5; DB 4; Length 469;

Best Local Similarity 20.7%; Pred. NO. 0.0047;

Matches 71; Conservative 61; Mismatches 118; Indels 93; Gaps 15;

QY 83 PECTAKVCAHAYVCGNDVYVNLQNTNL---QFNNNKYILI--QLLEDDAQRNFSVWNR 137  
DB 47 PQCRIVQVGRKRTII--NKLFFDLAQEEENYLDREFLKNELDNVRAQLSQDKKEDRS---- 100  
QY 138 WGRVCKMGQHSILVACSGNLNKAKEIFQKFLDKTKNNNEDRE---KPEKVPKGYDML-- 191  
DB 101 -----QVIIDTLRDTLEERNATVVSLOQALGKAEMLCS 133  
QY 192 ---QMDYATNTQDEETKKEES--LKSPKPEESQDLRVQELIKLINCVOAMEEMMEM 245  
DB 134 TLKKQMKYLEQQODETKQAQAEAGRLSRKMKMTQEIILLQS-----OLPEVEEMIRDM 187  
QY 246 KYNTKAPLKLTVAQIKAGYQSLAKIEDCIRAGOHGRALMEACNEFFYTRIPHDGLRTP 305  
DB 188 -----VGGSASVEQLAVYCVSLKKEYENLKEAR-----KASGEVADKLKRDLFSSRS 234  
QY 306 PLIRTKELSE-KIOLLEALGDIETAIKLVKTELQSPHEPLDQHYRNHLHLCARPLDHESY 364  
DB 235 KLOTVYSELDAQLELKSQKDLQADK-----EIMSLKKLTMLQETLN--LPPVASET- 287  
QY 365 EFKVISQYLSQSTHAPTHSDYTMTLDLDFEVEKDGKEAFREDL 407  
DB 288 ----VDRLVLESAPV-----EVLNKLRRPSFRDDI 314

RESULT 5

US-08-044-618-6

; Sequence 6, Application US/08044618

; Patent No. 5449605

GENERAL INFORMATION:  
APPLICANT: SMULSON, MARK  
TITLE OF INVENTION: METHOD OR DETECTING A PREDISPOSITION TO  
CANCER BY THE USED OF RESTRICTION FRAGMENT LENGTH  
POLYMORPHISM OF THE GENE FOR THE HUMAN POLY (ADP-RIBOSE)  
POLYMERASE  
TITLE OF INVENTION: POLYMERASE  
NUMBER OF SEQUENCES: 7  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Suite 300  
CITY: Washington  
STATE: D.C.  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/044,618  
FILING DATE: 19930406  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/07/257,696  
FILING DATE: 14-OCT-1988  
ATTORNEY/AGENT INFORMATION:  
NAME: FOX, SAMUEL L  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 0654, 0490001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)466-0800  
TELEFAX: (202)833-8716  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 163 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-044-618-6

Query Match 4.0%; Score 119; DB 1; Length 163;  
Best Local Similarity 28.3%; Pred. No. 0.0012;  
Matches 41; Conservative 27; Mismatches 57; Indels 20; Gaps 6;  
QY 29 APEDSSPAKTRCORQESKMPVAGGKANKDRYEDKODESVKALLKGPVDECTAK 88  
DB 15 APRGKGAVALSKSKGQV-----KEEGINKSEKPMK-LTLKGGAAVDPD-SGL 60  
QY 89 VGKAHVYCEGNDVYDMLNQTNLQFNKNKYLIQLLEDDAQRNFSVVMWGRVKG-MGQH 147  
DB 61 EBSAHVLEKGGKVSATLSLVVYKGTNSYYKLLKDDKESRHWIFKSWDRVGTVIGSN 120  
QY 148 SLVACSGNLNKAKEI--FQKFLDK 170  
DB 121 KL---EQMLSKEDTIEHFMLYEEK 142

RESULT 6  
US-08-968-751-2  
Sequence 2, Application US/08968751  
Patent No. 5948643  
GENERAL INFORMATION:  
APPLICANT: Rubinfeld, Bonnie  
APPLICANT: Polakis, Paul G.  
APPLICANT: Ligenfelter, Carol  
APPLICANT: Vuong, Terilyn T.  
TITLE OF INVENTION: MODULATORS OF BRCA1 ACTIVITY  
NUMBER OF SEQUENCES: 6  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: ONIX Pharmaceuticals, Inc.  
STREET: 3031 Research Drive  
CITY: Richmond

STATE: CA  
COUNTRY: USA  
ZIP: 94806  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/968,751  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Giotta, Gregory  
REGISTRATION NUMBER: 32,028  
REFERENCE/DOCKET NUMBER: ONYX1024 GG  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (510) 262-8710  
TELEFAX: (510) 222-9758  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 469 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-968-751-2  
Query Match 3.9%; Score 116.5; DB 2; Length 469;  
Best Local Similarity 20.7%; Pred. No. 0.011;  
Matches 71; Conservative 60; Mismatches 119; Indels 93; Gaps 15;  
QY 83 PECTAKYGAHVYCEGNDVYDMLNQTNL---OPNNKYYLI--QLLEDDAQRNFSVVMR 137  
DB 47 PQCRIQVGRKRTII--NKLFFDLAQEEENVLDLAEFLKKNELDNVRALQSOKDKEKRS---- 100  
QY 138 WGRVGMGQHSVLVACSGNLNKAKEIFOKKFLDKTKNNWEDRE----RFEKVPKGYDML-- 191  
DB 101 -----QVIIDTLRDTLEERNATVSVLQQAALKAEMLS 133  
QY 192 -----QMDYATNTQDEETKKEES--LKSPLKPESQLDRVQELIKLICNVQAMEMMEM 245  
DB 134 TLKQKMYLEQOQDETQKQAEARLRSMKMTBQIELLQS-----QRPEVEEMIRDM 187  
QY 246 KYNTKKAPLKLTVAQIKAGVQSLKKTEDCIRAGOHGRALMEACNEFYTRIPHDFGLRTP 305  
DB 188 -----GVQSQSAVEQLAVYCVSLKKEYENLKEAR-----KASGEVADKLKDLFSSRS 234  
QY 306 PLIRTKELSE-KIQLLEALGDIETIAIKLVKTELQSPHPLDQHYRNHLHCLALRPLDHESY 364  
DB 235 KLOTVYSELDOAKLELKSADKQADK----EIMSLKKKLTMLQETLN--LPPVASET- 287  
QY 365 EFKVISQYLOSTHAPTHSDYTWTLTLLDFEYKDKGEKAEFREDL 407  
DB 288 ----VDRVLVLESPAPV-----EYNLKLRRPSFRDDI 314  
RESULT 7  
US-08-609-230A-9  
Sequence 9, Application US/08609230A  
Patent No. 5866333  
GENERAL INFORMATION:  
APPLICANT: Innerarity, Thomas L.  
APPLICANT: Qian, Xiaobing  
APPLICANT: Yamanaka, Shinya  
TITLE OF INVENTION: Screening Methods to Detect mRNA Targets  
TITLE OF INVENTION: of Editing Enzymes  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: California



Db 293 WLEYESSFGSEEMQKVH-----ALRLASHLNLAMCHL-KLQAFSAATESCNKALELDSNN 346  
QY 250 KKAPLGK----LTVAQIKAGYQSLKTKEDCIRAGOHGRALMEACNEFYTRIPHDFGLRTP 305  
Db 347 EKLFRGEAHLAVNDFDLARDFQKVLQYPSNKAATOJAVCQ-----392  
QY 306 PLIRTOKEKSEKIQ-----LEALGDIEIAIKLVKTELOSPEHPLDQHYRNLHLCALRPLDH 361  
Db 393 ---RTRQLAREKKLYANMERLAEAE---HKVKAEEVAAGDHTDAERKSLPRVWPMDT 446  
QY 362 ESYEFKVISQYLOSTHAPTHS 382  
Db 447 K-----MQSLPTTHPHS 460  
  
RESULT 9  
US-08-687-080-51  
; Sequence 51, Application US/08687080  
; Patent No. 5965427  
; GENERAL INFORMATION:  
; APPLICANT: Gregory Dolganov  
; TITLE OF INVENTION: Human RAD50 Gene and Methods of Use Thereof  
; NUMBER OF SEQUENCES: 175  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dehlinger & Associates  
; STREET: 350 Cambridge Avenue, Suite 250  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/687,080  
; FILING DATE: 17-JUL-1996  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 08/592,126  
; FILING DATE: 26-JAN-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sholtz, Charles K.  
; REGISTRATION NUMBER: 38,615  
; REFERENCE/DOCKET NUMBER: 4600-0111.30  
; TELEPHONE: (415) 324-0880  
; TELEFAX: (415) 324-0960  
; INFORMATION FOR SEQ ID NO: 51:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1312 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHEICAL: NO  
; ANTI-SENSE: NO  
; ORIGINAL SOURCE:  
; INDIVIDUAL ISOLATE: TRANS. OF RAD50 cDNA (SEQ. 54), NT.  
; INDIVIDUAL ISOLATE: 389 TO 4324  
US-08-687-080-51  
  
Query Match 3.8%; Score 112.5; DB 2; Length 1312;  
Best Local Similarity 21.3%; Pred. No. 0.15;  
Matches 97; Conservative 71; Mismatches 150; Indels 137; Gaps 24;  
  
QY 5 RRSTGGGR-----ARALNESKRVN-----NGNT-----APEDSSPAK 37  
Db 433 RDKRTGRIIEIKSEILSKSRQKLNKVKYELQEGSSRIEILDOELKKAEREUSKAE 492  
QY 38 KTRRCORQESKMPVAGGKANKORTEDKODESVKAL-----LLKGKAPVDPEC 85

Db 493 KNSNVETLKMEVISLQNEKADLDRILRLKLDQEMEQLNHHHTTTRTQEMMLTKDKADK- EQ 551  
QY 86 TAKVGAHYVCENGDYDVMNLQTNQFNNNKYLLIQLLEDDAQRNFSVMMRWGRVCKMG 145  
Db 552 IRKIKSRH-----SDELTSLG-----YFPNKKO-----LED-----WL-----580  
QY 146 QHSLVACSGNLKAKKEIFQKFKDKTKNNWEDREKFEKVPKYDMLQMDYATNTQDEET 205  
Db 581 -HS-----KSKEINQ-----TRDLAKL--NKELASSEQKNHINNELK 616  
QY 206 KKEESLKPLKPSQLDLRVQELIKLICNVQAMEEMMKYNTKRAPLGLKLVQAKAG 265  
Db 617 RKEEQSLSS-----YEDKLFVCGSQDFESDLRLKEEIKSEKSKQRAMLAGATAV 665  
QY 266 Y-QSLKKIED-----CIRAGOHGRALMEACNEFYTRIPHDFGLRTP--LIRTOKEK 314  
Db 666 YSQFITQLTDENOSCCPVCORVFQTEALQEVISDLQSK-----LRLAPDKLKSTESEL 719  
QY 315 SEKIQLL-EALGDIEIAIKLVKTELOSPEHP-LDQHYRNLHLCALRPLDHESYEFKVISQY 372  
Db 720 KKEKREDEMLGLVPMRQSI--DLKEKEIPELRLNKLQNVNRDIOQLKNDIEE----QET 773  
QY 373 LOSTHAPTHSD-----YTMLLDLFEVE-KDGEKE 401  
Db 774 LLGTIMPEESAKVCLTDVTIMEFQEMELKDVERK 808  
  
RESULT 10  
US-08-764-100-27  
; Sequence 27, Application US/08764100  
; Patent No. 5773700  
; GENERAL INFORMATION:  
; APPLICANT: van Grinsven J., Martinus Q.  
; APPLICANT: De Haan, Petrus T.  
; APPLICANT: Gielen L., Johannes J.  
; APPLICANT: Peters, Dirk  
; APPLICANT: Goldbach, Robert W.  
; TITLE OF INVENTION: Improvements in or Relating to Organic  
; TITLE OF INVENTION: Compounds  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sandoz Agro, Inc  
; STREET: 975 California Avenue  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/764,100  
; FILING DATE: 06-DEC-1996  
; CLASSIFICATION: 800  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/214,064  
; FILING DATE:  
; APPLICATION NUMBER: US 08/032,235  
; FILING DATE: 17-MAR-1993  
; APPLICATION NUMBER: GB 9206016.9  
; FILING DATE: 19-MAR-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 5773700ris, Allen E.  
; REGISTRATION NUMBER: 34,490  
; REFERENCE/DOCKET NUMBER: 137-1061  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 354-3592  
; TELEFAX: (415) 857-1125  
; INFORMATION FOR SEQ ID NO: 27:  
; SEQUENCE CHARACTERISTICS:







Db 442 DLNNDENLMPEDQKLPVKKLDSKKKPKYSEVEKLDIKFSNNVNGELSPDKSSYKD 501  
QY 365 ---EFKVISO--YLOSTHAPHSYDTWTLLDLFEVEKDGKEAPRE----- 405  
Db 502 IDSKEETVKNQVNLQKTKPQVKDQVTSLSNEDLTMTSSSPVLEVIDPITNLGLQLI 561  
QY 406 DLHNRMLLWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKIYFADMSKSNYCFASR 465  
Db 562 DLNTGVRL----KSTQGIQRYGI-----YERKDLVVKMDSKAKLQIILDK 606  
QY 466 LKNTGLLLSEVALQCCNELLEANPKABGLLQGHSTKG--LGKMAPSSAHFVTLNGSTV 523  
Db 607 LENLKVVSSENFENKSSSLYVDSKMLVAVRDKDSSNDWLAKFSKP----- 654  
QY 524 PLGPASDTGILNPDGYTLNNEYIVYVNPQVVMRYL-LKVQFNFL 567  
Db 655 -----NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 14  
PCT-US92-05539-2  
; Sequence 2, Application PC/TUS9205539  
; GENERAL INFORMATION:  
; APPLICANT: Lefebvre, Rance B.  
; APPLICANT: Peng, Guey-Chen  
; TITLE OF INVENTION: Methods and Compositions for Diagnosing  
; TITLE OF INVENTION: Lyme Disease  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: James M. Heslin  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: USA  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/05539  
; FILING DATE: 19920629  
; CLASSIFICATION: 424  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Heslin, James M.  
; REGISTRATION NUMBER: 29,541  
; REFERENCE/DOCKET NUMBER: 2307U-330/UC 90-017-1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 415-326-2400  
; TELEFAX: 415-326-2422  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 700 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US92-05539-2

Query Match 3.7%; Score 110.5; DB 5; Length 700;  
Best Local Similarity 20.08; Pred. No. 0.084;  
Matches 11; Conservative 90; Mismatches 227; Indels 151; Gaps 26;  
QY 49 KMPVAGGKANDRTEDK--QDESVKALLKKGKAPVD--PECTAKVGAHVYCEGNDVYDV 104  
Db 189 KKDILSGNIESDIDISLTVTKVVAALLSENAAGVNFARDITDQETHK--AQDQKIDI 246  
QY 105 MLNQTNLQFNNNKYLIQLLEDQAQRNFVVMWRGVRGKMGQHSILVACSGNLNKAETQ 164  
Db 247 ELDNH-ESDSNITETIENLRD-----QLERATDEH 277

QY 165 KKFL----DKTKNWNEDREK----PEKVPKGYDMLQ--MDYATNT-----QDEETKKE 208  
Db 278 KKETESVDAKKKQKEELDKKAINLDKAAQKLDSEADNLVQRTVTRKIQEDINEIKE 337  
QY 209 ESLKSP---LKPESQDLRVOELIKLIGNVQAMEMMYNTKKAPLKITVAQIKAG 265  
Db 338 KNLKPKGVDVSPKVKQLQIKE-----SLEDLQBLKETGDENQKRETEK--QIEIKKS 389  
QY 266 YOSLKIKEDICRAGOHGRALMEACNEFYTRIPHDPLGLTPPLIRTPQKELSEKIQLEALG 325  
Db 390 DEKLKSKD-DKASDQKAL-DLDRELNSK-----ASSKEKSKAKEBEITKRSQKSLG 441  
QY 326 DIETAIKLVKTELAS-PE-----HPLDQHYR--NLHCALRPLDHESY-- 364  
Db 442 DLNNDENLMPEDQKLPVKKLDSKKKPKYSEVEKLDIKFSNNVNGELSPDKSSYKD 501  
QY 365 ---EFKVISO--YLOSTHAPHSYDTWTLLDLFEVEKDGKEAPRE----- 405  
Db 502 IDSKEETVKNQVNLQKTKPQVKDQVTSLSNEDLTMTSSSPVLEVIDPITNLGLQLI 561  
QY 406 DLHNRMLLWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKIYFADMSKSNYCFASR 465  
Db 562 DLNTGVSL----KSTQGIQRYGI-----YERKDLVVKMDSKAKLQIILDK 606  
QY 466 LKNTGLLLSEVALQCCNELLEANPKABGLLQGHSTKG--LGKMAPSSAHFVTLNGSTV 523  
Db 607 LENLKVVSSENFENKSSSLYVDSKMLVAVRDKDSSNDWLAKFSKP----- 654  
QY 524 PLGPASDTGILNPDGYTLNNEYIVYVNPQVVMRYL-LKVQFNFL 567  
Db 655 -----NLDEFILSENKIMPFTSFSVRKNFIYLQDEFKSL 688

RESULT 15  
US-08-990-140-4  
; Sequence 4, Application US/08990140A  
; Patent No. 6093795  
; GENERAL INFORMATION:  
; APPLICANT: Olsen, Henrik S.  
; APPLICANT: Ruben, Steven M.  
; APPLICANT: Sonenberg, Nahum  
; APPLICANT: Methot, Nathalie  
; APPLICANT: Rom, Eran  
; TITLE OF INVENTION: Human Pritl-like Subunit Protein (hPrt1) and Human  
; TITLE OF INVENTION: eIF4Gf-like Protein (p97) Genes  
; FILE REFERENCE: 1488.0700001  
; CURRENT APPLICATION NUMBER: US/08/990,140A  
; CURRENT FILING DATE: 1997-12-12  
; EARLIER APPLICATION NUMBER: US 60/033,151  
; EARLIER FILING DATE: 1996-12-13  
; NUMBER OF SEQ ID NOS: 13  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 907  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-08-990-140-4  
Query Match 3.7%; Score 110.5; DB 3; Length 907;  
Best Local Similarity 20.5%; Pred. No. 0.13;  
Matches 127; Conservative 87; Mismatches 196; Indels 211; Gaps 33;  
QY 2 AARRRSTGGGRALNPE--SKRVNN-----GNTAPEDSS---PAKTRRCORQESKKMPV 52  
Db 10 ASRESASSGGGSGRAGPOHYPKTAGNSEFLGKTPGQNAQKWIPARSTRDONS-----A 63  
QY 53 AGGKANKRDTEDKQDESVKALLKKGKAPVDPECTAK-----VG-KAHVYCEGNDVYDV 104  
Db 64 ANNSANERKHDAIFRKVRGILNK----LTPEKFDKLCLELLNVGVEKSLILKG-----V 114  
QY 105 MLNQTNLQFNNNKY-----YLIQLLEDQAQRNFVVMWRGVRGKMGQHSILVACSGNLNKA 159

Db	115	ILLIVDKALBEPKYSSIIYAQILCLRLAEDAP-NFDGPAEQCPQ-----KQ	159
Qy	160	KEIFQKKFLDKTKNNWNEDREKFVKPGKYDMLQMDYATNQDEETTKBESLKSPLKPES	219
Db	160	STTFRRLLISKLQDEFENRTRNDVV--YD-----KRE-----NPLLPEE	196
Qy	220	QLDLRVQELTKLINCVOAMEMMEMKYNTKKAPLKGKLFVAQIKAGYQSLKIEDCIRAG	279
Db	197	E-EQRAIAKIMLGNIKFGE-----LGLK-----DLIH-----224	
Qy	280	QHGRALMEACNEFVTRIPHDGFLRPLIRIQKSELSEKQLLEALGDIEIAIKLVKLTQLQ	339
Db	225	--ESILHKC-----IKTLEKKRVLQKDMGEDLELCQIMRTVCP	263
Qy	340	SPEHP-----LDQHYRNHLCAPLRDHESYEFKVISQYLOSTHAPHSDYTWTLDLDFEV	394
Db	264	RLDHERAKSLMDQYFARM--CSL-----MLSKELPARIRF--LLQDTVEL	304
Qy	395	EKDG--EKEAFREDLHNRMLLWHGSRMSNV-----GI-----LSHLR-----431	
Db	305	REHHWVRKAF-----LDNGPKYINQIQADVAKDVGVFIPMAAGMSRSDFELEG	355
Qy	432	IAPPE-----APITGYMFGKI-----YFADMSSKSNYCFASRLKNTGLLLLL	475
Db	356	FMPPRMKMDRDLGGLADMFGMPGSGIGTGPQVIQDRFSPTMGRRHSNQLFNHGHHIM	415
Qy	476	EVALGQCNELLEANPRAEGLQCKHS--TKGLGKMWSSAHFVTNGSTVPLGPA--SDTCI	533
Db	416	PPTQSQFGMGCGFMKSSQGLSOLYIHQNSQGL-----LSOLOQSGKMDPPRFSSKKGO	466
Qy	534	LNPDGYTLNINEYIVYNPNOV	554
Db	467	LNADEISLRPAQSGFLMNKNQV	487

Search completed: August 29, 2002, 07:58:08  
Job time: 285 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:57:23 ; Search time 36.19 Seconds  
(without alignments)  
609.841 Million cell updates/sec

Title: US-09-701-586B-2

Perfect score: 2998

Sequence: 1 MAARRRRSTGGGRARALNES.....PNQVRMYLLKVFNFQLW 570

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2981.5	99.4	583	1 PPO2_HUMAN	Q9ug55 homo sapien
2	2565.5	85.6	559	1 PPO2_MOUSE	Q88554 mus musculu
3	1180	39.4	637	1 PPO2_ARATH	Q11207 arabidopsis
4	1027	34.3	1011	1 PPO2_CHICK	P26446 gallus gall
5	1022	34.1	1012	1 PPO2_CRIGR	Q9rl52 cricetus
6	1017	33.9	1012	1 PPO2_MOUSE	P11103 mus musculu
7	1016	33.9	1013	1 PPO2_RAT	P27008 rattus norv
8	1012.5	33.8	1013	1 PPO2_HUMAN	P09874 homo sapien
9	1009	33.7	1015	1 PPO2_BOVIN	P18493 bos taurus
10	984.5	32.8	998	1 PPO2_XENLA	P31669 xenopus lae
11	979.5	32.7	994	1 PPO2_DROME	P35875 drosophila
12	964	32.2	996	1 PPO2_SARPE	Q11208 sarcophaga
13	720	24.0	533	1 PPO3_HUMAN	Q9y6f1 homo sapien
14	526	17.5	538	1 YON4_CAEEL	Q09525 caenorhabdi
15	331.5	11.1	135	1 PPO2_ONCMA	Q08824 oncorhynchu
16	304.5	10.2	1724	1 PPOV_HUMAN	Q9ukk3 homo sapien
17	133	4.4	805	1 SEC6_YEAST	P32844 saccharomyc
18	130.5	4.4	1005	1 RA50_METJA	Q58718 methanococc
19	128	4.3	2230	1 GOG4_HUMAN	Q13439 homo sapien
20	126	4.2	880	1 RA50_PYRAB	Q9uzc8 pyrococcus
21	126	4.2	2058	1 MY10_HUMAN	Q9hd67 homo sapien
22	122	4.1	1938	1 MYH9_HUMAN	Q9ukx3 homo sapien
23	121.5	4.1	724	1 HMR_HUMAN	Q05530 homo sapien
24	121.5	4.1	794	1 HMR_MOUSE	Q00547 mus musculu
25	119.5	4.0	5430	1 ACF7_HUMAN	Q9upn3 homo sapien
26	118.5	4.0	1325	1 G160_MOUSE	P55937 mus musculu
27	117.5	3.9	1940	1 KINH_CHICK	P02565 gallus gall
28	117	3.9	963	1 MYH1_HUMAN	P33176 homo sapien
29	116.5	3.9	886	1 RA50_ARCFU	Q09230 archaeoglob
30	116	3.9	504	1 YLF3_CAEEL	Q03573 caenorhabdi
31	115.5	3.9	540	1 YKZ6_YEAST	P36112 saccharomyc
32	115.5	3.9	600	1 PFR1_TRYBB	P22225 trypanosoma
33	115	3.8	1790	1 USO1_YEAST	P25386 saccharomyc

34	115	3.8	1846	1 MY5B_RAT	P70569 rattus norv
35	114.5	3.8	982	1 P115_MYCGE	P47540 mycoplasma
36	114.5	3.8	1959	1 MYH9_CHICK	P14105 gallus gall
37	114	3.8	734	1 YHU9_YEAST	P32900 saccharomyc
38	114	3.8	1230	1 SMC3_YEAST	P47037 saccharomyc
39	113	3.8	963	1 KINH_MOUSE	Q61768 mus musculu
40	112.5	3.8	1650	1 BP28_CAEEL	Q23495 caenorhabdi
41	112.5	3.8	1978	1 MYH8_CHICK	P10587 gallus gall
42	112	3.7	770	1 KLPA_EMENI	P28739 emeritocella
43	111.5	3.7	727	1 MFPL_ARATH	Q91w85 arabidopsis
44	111.5	3.7	2035	1 EVPL_MOUSE	Q94952 mus musculu
45	111	3.7	876	1 PRO2_YEAST	P20095 saccharomyc

ALIGNMENTS

RESULT	1
ID	PPO2_HUMAN
AC	Q9UGN5; Q9Y6C8; Q9NUV2; Q9UMR4; STANDARD; PRT; 583 AA.
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	01-MAR-2002 (Rel. 41, Last annotation update)
DE	Poly (ADP-ribose) polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-riboseyltransferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2) (hPARP-2).
DE	(hPARP-2).
GN	ADPRT12 OR PARP2 OR ADPRT2..
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A. (ISOFORM 2).
RC	TISSUE=Fetal brain;
RX	MEDLINE=99292755; PubMed=10364231;
RA	Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P., Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RA	"PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase.";
RT	J. Biol. Chem. 274:17860-17868(1999).
RL	[2]
RN	SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RP	TISSUE=Fetal brain;
RC	MEDLINE=99263509; PubMed=10329013;
RX	Johansson M.;
RA	"A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA cloning of two novel poly(ADP-ribose) polymerase homologues.";
RT	Genomics 57:442-445(1999).
RN	[3]
RP	SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC	TISSUE=Fibroblast;
RX	MEDLINE=99268466; PubMed=10338144;
RA	Berghammer H., Ebner M., Marksteiner R., Auer B.;
RA	"pADPRT-2: a novel mammalian polymerizing (ADP-riboseyl)transferase gene related to truncated pADPRT homologues in plants and Caenorhabditis elegans.";
RT	FEBS Lett. 449:259-263(1999).
RN	[4]
RP	SEQUENCE FROM N.A. (ISOFORM 1).
RC	TISSUE=Placenta;
RA	Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y., Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y., Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Kanehori K.;
RA	"NEDO human cDNA sequencing project.";
RT	Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.
CC	-1- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY. SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (By similarity).
CC	-1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor = nicotinamide + {ADP-D-riboseyl}(N+1)-acceptor.

CC -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).  
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be  
CC produced by alternative splicing.  
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN  
CC THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO  
CC DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;  
CC LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND  
CC THYMUS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; AJ236912; CAB65088.1; .  
CC DR EMBL; AF085734; AAD29857.1; ALT INIT.  
CC DR EMBL; AJ236876; CAB41505.2; ALT INIT.  
CC DR EMBL; AK001980; BAA92017.1; ALT\_TERM.  
CC DR HSSP; P26446; I26.  
CC InterPro: IPR001290; PARP.  
CC DR InterPro: IPR004102; PARP\_reg.  
CC DR Pfam; PF00644; PARP; 1.  
CC DR Pfam; PF02877; PARP\_reg; 1.  
CC KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Alternative splicing.  
FT DNA\_BIND 1 88 POTENTIAL.  
FT DOMAIN 86 583 NAD-BINDING (BY SIMILARITY).  
FT DOMAIN 4 7 NUCLEAR LOCALIZATION SIGNAL 1ST PART  
FT DOMAIN 35 40 NUCLEAR LOCALIZATION SIGNAL 2ND PART  
FT DOMAIN (POTENTIAL).  
FT VARSPLIC 68 80 MISSING (IN ISOFORM 2).  
FT CONFLICT 447 447 P -> H (IN REF. 2).  
FT CONFLICT 481 481 N -> H (IN REF. 4).  
FT SEQUENCE 583 AA; 66205 MW; 5B7AE8AE531836AF CRC64;

Query Match 99.4%; Score 2981.5; DB 1; Length 583;  
Best Local Similarity 97.8%; Pred. No. 1.4e-187;  
Matches 570; Conservative 0; Mismatches 0; Indels 13; Gaps 1;  
QY 1 MAARRRSTGGGRARALNESKRYNNGTAPEDSSPAKTRRCORQESKMPVAGGKAND 60  
DB 1 MAARRRSTGGGRARALNESKRYNNGTAPEDSSPAKTRRCORQESKMPVAGGKAND 60  
QY 61 RTEDKOD-----ESVKALLLKGKAPVDPCTAKVGHVYCGNDVDVMLN 107  
DB 61 RTEDKODGMPGRSWASKRVSSEVKALLLKGKAPVDPCTAKVGHVYCGNDVDVMLN 120  
QY 108 QTNLQFNNNKYIQLLEDAAQRNFSVMWRGVRGKMGHSLVACSGNLNKAKEIFQKF 167  
DB 121 QTNLQFNNNKYIQLLEDAAQRNFSVMWRGVRGKMGHSLVACSGNLNKAKEIFQKF 180  
QY 168 LDKTKNWDREKFEKVPKGYDMLQMDYATNTQDEETKKEESLKSPKPSOLDLRVQE 227  
DB 181 LDKTKNWDREKFEKVPKGYDMLQMDYATNTQDEETKKEESLKSPKPSOLDLRVQE 240  
QY 228 LKILCNVQAEEMEMKYNKTKAPLGLTVAQIRAGVQSLKKEIECDIRAGQGRALME 287  
DB 241 LKILCNVQAEEMEMKYNKTKAPLGLTVAQIRAGVQSLKKEIECDIRAGQGRALME 300  
QY 288 ACNEFYTRIPDHGLRTPPLIRTOKESEKIQLEALGDIETAIKLVKTELQSPHPLDQ 347  
DB 301 ACNEFYTRIPDHGLRTPPLIRTOKESEKIQLEALGDIETAIKLVKTELQSPHPLDQ 360  
QY 348 HYRNLCALRPLDHESEKVFVISOYLOSTHAPTHSDYTWTLTLLDFEVEKDKGEAREDL 407  
DB 361 HYRNLCALRPLDHESEKVFVISOYLOSTHAPTHSDYTWTLTLLDFEVEKDKGEAREDL 420

QY 408 HNRMLLWHGSRMSNWVYILSHGLRIAPPEAPITGYMFGKIYFADMSKSNYCFASRLK 467  
DB 421 HNRMLLWHGSRMSNWVYILSHGLRIAPPEAPITGYMFGKIYFADMSKSNYCFASRLK 480  
QY 468 NTGLLLSEVALGQCNELEANPKAEGLLQGHKSTKGLGKMAPSSAHFVTLNGSTVPLGP 527  
DB 481 NTGLLLSEVALGQCNELEANPKAEGLLQGHKSTKGLGKMAPSSAHFVTLNGSTVPLGP 540  
QY 528 ASDTGILNPDGYTLNEXIVYVNPNOVMRYLLKVKQFNFLQLW 570  
DB 541 ASDTGILNPDGYTLNEXIVYVNPNOVMRYLLKVKQFNFLQLW 583  
RESULT 2  
PPO2\_MOUSE  
ID PPO2\_MOUSE STANDARD; PRT; 559 AA.  
AC O88554; Q99N29;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-  
DE ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2)  
DE (mPARP-2).  
GN ADPRTL2 OR PARP2 OR ADPRT2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Embryo;  
RX MEDLINE=99292755; PubMed=10364231;  
RA Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,  
RA Muller S., Heger T., Menissier-de Murcia J., de Murcia G.M.;  
RT "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)  
RT polymerase.";  
RL J. Biol. Chem. 274:17860-17868(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SV;  
RX MEDLINE=21179160; PubMed=11133988;  
RA Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,  
RA Niedergang C.P.;  
RT "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2  
RT (PARP-2) gene to the gene for RNase P RNA.";  
RL J. Biol. Chem. 276:11092-11099(2001).  
RN [3]  
RP SEQUENCE OF 9-559 FROM N.A.  
RC STRAIN=129/SV X C57BL/6;  
RX MEDLINE=99268466; PubMed=10338144;  
RA Berghammer H., Ebner M., Marksteiner R., Auer B.;  
RT "ADPRT-2: a novel mammalian polymerizing(ADP-ribose)transferase gene  
RT related to truncated PADPRT homologues in plants and Caenorhabditis  
RT elegans.";  
RL FEBS Lett. 449:259-263(1999).  
CC -!- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY.  
CC -!- SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose](N)-acceptor =  
CC nicotinamide + [ADP-D-ribose](N+1)-acceptor.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- TISSUE SPECIFICITY: Widely expressed; the highest levels were in  
CC testis followed by ovary.  
CC -!- INDUCTION: By high levels of DNA-damaging agents.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC







[illegible]

```
QY 87 AKVGRHVCYCEGNDVYVLMNQTINQFNNNKYYLIQLLEDDAQRNFSSVMMRWGRYCKMGQ 146
DB 536 GLEHSAHVLEKGGKVSATLGLVDIVKGTNSYKQLLEDDKESRWIFRSRGRVGTG 595
QY 147 HSLVACSGNLNKAKEIFQKFKLDKTKNNWEDREKFEKVPKGYDMLQMDYATNTQDEEFTK 206
DB 596 SNKLEQMPKEDAVEHFMKLYBEKTNWHSK-NETKPKKFPLEIDYG---QDEEAVK 651
QY 207 KEESLKSPLKP--ESOLDLRVQELIKLNCVQAEEMMEMMYNTKAPLGLKLTVAQIKA 264
DB 652 -----KLATPKGTGSKLPAVQDELGMIFDVESMKALVEYIDLQKPLGLSKSRQIOA 706
QY 265 GYQSLKKIEDCTRAGHGRALMEACNEFYTRIPHDPLRTPLIRTKQKELSEKIOLEAL 324
DB 707 AVSILSEVQAVSQGSDSQILDLSNRFYTLPHDEGMKPPLLNADSVQAKVEMLDNL 766
QY 325 GDIEIAIKLVK-TELQSPHPLDQHYRNHLCALPLDHSYEFKVISQYLSTHAPTHSD 383
DB 767 LDIEVAYSLRGSGDDSDPDIDVNYEKLKTDIKVVDRODSEAEVIRKVKVNTHTATHNA 826
QY 384 YMTLLDLFEVEKCEKFAFR--EDLHNRMLLWHSRMSNWGILSHGLRIAPPEAPITG 441
DB 827 YDLEVMDFIEKIEREGESQRYKFPKQLHNRLLWHSRRTTFAGILISQGLRIAPPEAPVTG 886
QY 442 YMFQGIYFADMSKSSANCYCFASRLKNTGLLLSEVALGQCNEELLEAPNKAEGLLQGRHS 501
DB 887 YMFQGIYFADVMKSKSANCYCHTSQGDPIGLILLGEVALGNMYELKHASHISK-LPKGRHS 945
QY 502 TKGGLKMAFSSAHFVNLNGSTVPLGSPASTGILNPDGYT---LNYNEYIVYNNQVRY 558
DB 946 VKGLGRTTDPDSASITLEGVEVPLG---TGI--PSGVNDTCLLYNEYIVYDIAQVNLKY 999
QY 559 LLKVOQNF-LQLW 570
DB 1000 LUKLRFNFTSLW 1012

RESULT 6
PQOL_MOUSE
ID PQOL_MOUSE STANDARD; PRT; 1012 AA.
AC P11103; Q9JLX4; Q9VQ03;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE ADP-riboseyltransferase-1 (Poly[ADP-ribose] synthetase-1) (msPARP)
GS ADPRT OR ADPRT1 OR ADPRT.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=BXSB;
RX MEDLINE=89263780; PubMed=2498841;
RA Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;
RT "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT gene.";
RL Nucleic Acids Res. 17:3387-3401(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC STRAIN=129/SV X C57BL/6; TISSUE=Fibroblast;
RX MEDLINE=20270268; PubMed=10809783;
RA Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT with poly(ADP-ribose) polymerase activity independent of DNA strand
RT breaks.";
RL J. Biol. Chem. 275:15504-15511(2000).
RN [3]
RP KNOCK-OUT.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haldacher D., Jaeger S., Berghammer H.,
Kofler B., Schweiger M., Wagner E.F.;
"On the biological role of the nuclear polymerizing NAD(+): protein (ADP-ribose)yl transferase (ADPRT): ADPRT from Dictyostelium discoideum and inactivation of the ADPRT gene in the mouse.";
Biochimie 77:444-449(1995).
CC FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT
ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =
nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC SUBUNIT: HOMODIMER (Potential).
CC SUBCELLULAR LOCATION: Nuclear.
CC ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
short form/sPARP-1; may be produced by alternative initiation.
CC MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC SIMILARITY: BELONGS TO THE PARP FAMILY.
CC SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
EMBL; X14206; CA32421.1; -;
EMBL; AF126717; AAF61293.1; ALT_INIT.
PIR; S04200; S04200.
HSP; P26446; I426.
MGD; MGI:1340806; Adprt1.
InterPro; IPR001357; BRCT.
InterPro; IPR001290; PARP.
InterPro; IPR004102; PARP-reg.
InterPro; IPR001510; Znf-PARP.
Pfam; PF005533; BRCT; 1.
Pfam; PF006644; PARP; 1.
Pfam; PF02877; PARP-reg; 1.
Pfam; PF00645; zf-PARP; 2.
ProDom; PD004675; Znf-PARP; 2.
SMART; SM00292; BRCT; 1.
PROSITE; PS50172; BRCT; 1.
PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
PROSITE; PS00347; PARP_ZN_FINGER_2; 2.
PROSITE; PS00664; PARP_ZN_FINGER_2; 2.
Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.
INIT_MET 0 0 BY SIMILARITY.
CHAIN -1 1012 POLY [ADP-RIBOSE] POLYMERASE-1, LONG
ISOFORM.
CHAIN 521 1012 POLY [ADP-RIBOSE] POLYMERASE-1, SHORT
ISOFORM.
CHAIN 521 1012 FOR SHORT ISOFORM.
INIT_MET 521 521
DNA_BIND 1 371 AUTOMODIFICATION DOMAIN.
DOMAIN 372 522 BRCT.
DOMAIN 384 460 NAD-BINDING.
DOMAIN 523 1012 PARP-TYPE.
ZN_FING 20 55 PARP-TYPE.
ZN_FING 124 161 NUCLEAR LOCALIZATION SIGNAL 1ST PART.
DOMAIN 206 208 NUCLEAR LOCALIZATION SIGNAL 2ND PART.
DOMAIN 220 225 ADP-RIBOSYL[N] (POTENTIAL).
MOD_RES 406 406 ADP-RIBOSYL[N] (POTENTIAL).
MOD_RES 412 412 ADP-RIBOSYL[N] (POTENTIAL).
MOD_RES 434 434 ADP-RIBOSYL[N] (POTENTIAL).
MOD_RES 436 436 ADP-RIBOSYL[N] (POTENTIAL).
MOD_RES 443 443 ADP-RIBOSYL[N] (POTENTIAL).
```

FT	MOD_RES	444	444	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	447	447	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	455	455	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	483	483	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	487	487	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	490	490	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	511	511	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	512	512	ADP-RIBOSYL[N]	(POTENTIAL).
FT	MOD_RES	518	518	ADP-RIBOSYL[N]	(POTENTIAL).
FT	CONFLICT	590	590	L -> V (IN REF. 2).	
FT	CONFLICT	607	607	E -> D (IN REF. 2).	
FT	CONFLICT	611	611	Q -> H (IN REF. 2).	
FT	CONFLICT	628	628	N -> D (IN REF. 2).	
FT	CONFLICT	678	678	D -> E (IN REF. 2).	
FT	CONFLICT	702	702	R -> F (IN REF. 3).	
FT	CONFLICT	716	716	Q -> E (IN REF. 2).	
FT	CONFLICT	757	757	Q -> L (IN REF. 2).	
FT	CONFLICT	856	856	R -> F (IN REF. 3).	
FT	CONFLICT	981	981	A -> C (IN REF. 2).	
SQ	SEQUENCE	1012	AA; 112968	MW; 4354C3E5F01B9439	CRC64;
Query Match 33.9%; Score 1017; DB 1; Length 1012;					
Best Local Similarity 40.1%; Pred. No. 6.9e-59;					
Matches 219; Conservative 113; Mismatches 178; Indels 36; Gaps 12;					
Qy	31	EDSSPAKTRCORQESKKMPVAGGKANKRDKQDESVALLLKGPVDPCTAKVG	90		
Db	497	KSAPSKKSGCFKEG-----VNKSEKRMK-LTLKGGAADPD-SGLEH	539		
Qy	91	KAHYCEGNDVYDVLNQTNNKYLQLLEDDAQRNFSVMWRGKMGQHSVLV	150		
Db	540	SAHVLEKGGKVFSAFLGLVDIVKGTNSYKQLLEDDKESYWPFSRGLGTVIGSNKL	599		
Qy	151	ACSGNLNKAKEIFQKFKDKTKNNWEDREKEKVPKGYDMLQMDYATNQDEEFTKEES	210		
Db	600	EQMPSKEEAQFMKLYEKTGNAMHSK-NFTKYPKKEYPLEIDYG---QDEEAVK----	651		
Qy	211	LKSPKLP--ESOLDLRVQELIKLICNVQAEEMMEMMYNTKKAPLGKLTVAQIKAGYQS	268		
Db	652	-KLTVKPTKSKLPKPVQELVGMIFDVMKRAKALVEYEDIQKPLGKLSRRQQAAYS	710		
Qy	269	LKKIEDIRAGOHGRALMEACNEFYTRIPDFGLRTPLIRTQKELSEKIQLEALGIE	328		
Db	711	LSEVQPVQSQSSQILDSNRFYTLIPDFGKMKPPLLNADSVQAKVEMLDNLDDIE	770		
Qy	329	IAIKLVK-TELQSPHPDQHYRNHCAALRPLDHESYEFKVISQYLOSTHAPTHSDYPTM	387		
Db	771	VAYSLLRGSDSDSKDPIDVNYEKLTKTDIKVVDRESEAEVIRKVKVKNTHATTHNAYDLE	830		
Qy	388	LLDLFEVKGDEKEAFR--EDLHNRMLLWHSRMSNMWYGLSHGLRIAPPEAPITGYMFG	445		
Db	831	VIDIFKIEREGESQRYKFPFQRLHNRLLWHSRRTNFAGILSQGLRIAPPEAPVGYMFG	890		
Qy	446	KGIYFADSSKSANYCFASIKNTGTLLELSEVALGOCELEAKPKAGLQGHSTKGL	505		
Db	891	KGIYFADVMSKSNYCHTSQGDPTGLIMLGEVALGNMVELKHASHISK-LPKGHSYKGL	949		
Qy	506	GKMAPSSAHFVTLNGSTVPLGASDTGLNPDGYTLNNEYIVYNPNQVMRYLLKLVQFN	565		
Db	950	GKTPDPDSASITLEGVEVPLGTGIPSGV---NDTALLYNEIVYDIAQVNLKLLKLFN	1006		
Qy	566	F-LQLW	570		
Db	1007	FKTSLW	1012		

RESULT 7

PPOL_RAT	STANDARD;	PRT; 1013	AA.
AC	PPOL_RAT		
DT	01-AUG-1992 (Rel. 23, Created)		
DT	15-JUL-1998 (Rel. 36, Last sequence update)		

16-OCT-2001 (Rel. 40, Last annotation update)  
Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
ADPRT.  
Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Monocytes;  
RX MEDLINE=98046546; PubMed=9385436;  
RA Beneke S., Meyer R., Buerkle A.;  
RT "Isolation of cDNA encoding full-length rat (Rattus norvegicus) poly  
(ADP-ribose) polymerase.";  
RL Biochem. Mol. Biol. Int. 43:755-761(1997).  
RN [2]  
RP REVISION TO 811.  
RA Beneke S., Meyer R., Buerkle A.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE OF 1-11 FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;  
RX MEDLINE=92290013; PubMed=1601134;  
RA Potvin F., Thibodeau J., Kirkland J.B., Dandenault B.,  
R Duchaine C., Polier G.G.;  
RT "Structural analysis of the putative regulatory region of the rat  
gene encoding poly(ADP-ribose) polymerase.";  
RL FEBS Lett. 302:269-273(1992).  
RN [4]  
RP SEQUENCE OF 514-1013 FROM N.A.  
RC STRAIN=SPRAGUE-DAWLEY; TISSUE=Prostate;  
RX MEDLINE=90027702; PubMed=2508731;  
RA Thibodeau J., Gradwohl G., Dumas C., Clairoux-Moreau S., Brunet G.;  
RT "Cloning of rodent cDNA coding the poly(ADP-ribose) polymerase  
catalytic domain and analysis of mRNA levels during the cell cycle.";  
RL Biochem. Cell Biol. 67:653-660(1989).  
CC -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor -  
nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; U94340; AAC53544.1; -;  
DR EMBL; X65496; CAA46477.1; -;  
DR EMBL; X65497; CAA46478.1; ALT\_INIT.  
DR HSSP; P26446; 1A26.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00553; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.





Db 493 APRKSGAALSKKSGQV-----KEEGINKSEKRMK-LTLKGGAAVDPD-SGL 538  
QY 89 VGKAVHYCEGNDVYDMLNOTLNQNNKYYLQLEDDAORNFVSMRWGRVGMQGH 148  
Db 539 EHSADVLEKGGKVFATLGLVDIVKGTNSYKQLLEDKKNRYWIFRSWGRVGTIGSN 598  
QY 149 LVACSGNLNKAKEIFQKFLDKTKNNWEDREKFEKVPKGYDMLQMDYATNTQDEETKE 208  
Db 599 KLEQMPKSKEDAIEHPMKLYEEKTGNASHK-NFTKPKKFPLEIDYG---QDEAVKK- 653  
QY 209 ESLKSPKPEQSLDRVQELKLCINQVQAEEMEMMYNTKAPLGLKLTVAQIKAGTOS 268  
Db 634 --LTVNPGTKSLPKPVQDLKMIQFVDSMKAMVEYIDLQKMLGKLSKRQIQAAYSI 711  
QY 269 LKKIEDCTRAGHGRALMEACNEFYTRIPDFGLTPPLIRTKQELSEKIOLEALGDIE 328  
Db 712 LSEVQAVSQGSDSDILDSNRFYTLPHDFGKMKPPLNADSVQAKVEMLDNLIDIE 771  
QY 329 IAIKLVK-TELQSPHPLDQHYRNLCALRPLDHESYEFKVISQYLSQTHAPTHSDYTM 387  
Db 772 VAYSLRGSDSDSDIPDVNVEKLTIDIKVVDSDSEAEIIRKVVKNTHATTHAYDLE 831  
QY 388 LLDLEVEKGEKEAFR--EDLHNRMLLWHGSRMSNWYGLSHGLRIAPPEAPITGYMFG 445  
Db 832 VIDIFKIEREGECQRYKPKFKQLHNRLLWHGSRRTTFAGILSQGLRIAPPEAPVTYMFG 891  
QY 446 KGIYFADMSKSNYCFASRLKNTGLLLSEVALGOCNELLEANPKAELGQKHSKGL 505  
Db 892 KGIYFADMSKSNYCHTSQGDPIGLILLGEVALGNMYELKHASHK-LPKGKHSVGL 950  
QY 506 GKMAPSSAHFVTLNGSTVPLGPASDTGILNPDGYTLNVEYIVNPNQVMRYLLKLVQFN 565  
Db 951 GKTTDPSPANISLDGVDVPLGTGISSGV---NDTSLLYNEYIVYDIAQVNLKYLKLFN 1007  
QY 566 F-LQLW 570  
Db 1008 FKTSLW 1013

RESULT 9  
PPOL\_BOVIN STANDARD; PRT; 1015 AA.  
AC P18493; Q9TS00;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
GN ADPRT.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=90382673; PubMed=2119324;  
RA Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;  
RT "Cloning of a full-length cDNA encoding bovine thymus  
RT poly(ADP-ribose) synthetase: evolutionarily conserved segments and  
RT their potential functions.";  
RL Gene 90:249-254(1990).  
[2]  
RN SEQUENCE OF 647-714 AND 838-903 FROM N.A.  
RX MEDLINE=88151954; PubMed=2450019;  
RA Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,  
RA Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;  
RT "Depression in gene expression for poly(ADP-ribose) synthetase during  
RT the interferon-gamma-induced activation process of murine macrophage  
RT tumor cells.";  
RL Eur. J. Biochem. 171:571-575(1988).  
CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT

Query Match

33.7%; Score 1009; DB 1; Length 1015;





```
Query Match          32.8%; Score 984.5; DB 1; Length 998;
Best Local Similarity 39.1%; Pred. NO. 9e-57;
Matches 222; Conservative 109; Mismatches 196; Indels 41; Gaps 14;

QY 21 KRVNNGNTAPE-----DSSPAKTRRCQRCQESKKMPVAG-----GKANKRTEDEKQD 67
DB 444 KEVSGKSVQELLSOGFISSWGAETIKQBAVQTEKQPSGSGVACKSSGKVEEKGSKNSE 503
QY 68 ESKVALLKGRAPVDPECTAKVGRHVVCEGNDVYVDMNQTNIQFNKNNKYLILQLEDD 127
DB 504 KMK-LTVKGGAADPDESELE-DSCHVLETGKTFSATLGLVDTIRGTSNYYKQLLIEHD 561
QY 128 AQRFNSVMWRGVGK-MGQSHLVACSGNLNKAKEIFQKFLDKTKNNWDERKFKVPG 186
DB 562 RDSRWYFWSRGVGTGVSIGKLEEMSSK-EDATEHFLNLYQDKTGNNAWHS-PNFTKYPK 619
QY 187 KYDMLQMDYANTQDEETKEESLSKPLKPESQLDLRVQLIKLNCVQMEEMEMK 246
DB 620 KFYPLEIDYG---QBEDVYVK---LSVGAGTKSKLAKPVQELIKLIFDVESMKAMVEFE 673
QY 247 YNTKAPGLKLTVAQIKAGYQSLKKIEDCIRAGQHGRLMEACNEFYTRIPHDGFLRTPP 306
DB 674 IDLQKMPGLKSKRQIQSAYSILSOVQAVSESLSEARLLDLSNQFYTLIPHDGFMKKPP 733
QY 307 LIRTOKELSEKIQLEALGDIEIAIKLVKTELQSP-HPLDQHYRNHLHCLRLPLDHESYE 365
DB 734 LLNNLEYIQAKVQMLDNLDEIVAYSLLRGGADGDDPKIDVKYKIKTDIKVAKDSEE 793
QY 366 PKVTSYQLOSTHAPTHSYDTWTLLDLFEVKGDEKEAPR--EDLHNRMLLWHGSRMSNW 423
DB 794 SRICIDYVKNTHADPHNAYDLEVLIEIFKIDREGEYQRYKPKQLHNRQLLWHGSRNTNFA 853
QY 424 GILSHGLRIAPPEAPITGYMGKGIYFADMSKSKSANYCFASRLKNTGILLLLSEVALGOCN 483
DB 854 GILSGLRIAPPEAPVGYMGKGIYFADMSKSKSANYCHAMPGPSIGLILGEVALGNMH 913
QY 484 ELLEANPRAEGELGCKHSTKGLKMAPSSAHEFTLNGSTVPLGPA-----SDTGILNPDG 538
DB 914 E-LKAASQITKLPKHKSVKGLGRTPADPSATVQLDGVDPVPLGKGTSANISDTSLL---- 968
QY 539 YTLNNEYIVYNPQVRNRYLLKQVNF 566
DB 969 ----YNEIVYDIAQVNLKLLKLFNY 992

RESULT 11
PPOL_DROME          STANDARD;          PRT;          994 AA.
AC P35875; Q9W5S1;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE Poly (ADP-ribose) polymerase (EC 2.4.2.30) (PAP) (ADPRT) (NAD(+)) ADP-
DE ribosyltransferase [Poly(ADP-ribose) synthetase].
GN PAP OR CG17696/CG1718.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93234521; PubMed=8475096;
RA Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,
RA Sugimura T., Miwa M.;
RT "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:
RL leucine zipper in the auto-modification domain.";
RL Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).
RN [2]
RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND
RP TISSUE SPECIFICITY.
RC STRAIN=CANTON-S;
RX MEDLINE=98234380; PubMed=9565614;
```

```
RA Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;
RT "Genomic organization of Drosophila poly(ADP-ribose) polymerase and
RL distribution of its mRNA during development.";
RL J. Biol. Chem. 273:11881-11886(1998).
RN [3]
RP SEQUENCE FROM N.A. (LONG ISOFORM).
RC STRAIN=BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.G., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson G.L.G.,
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Balaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslcek A., Gong F., Gorrell J.H., Gu Z., Guan P., Glasser K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J., Yao Q.A.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhou X., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou S., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
CC -|- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY[ADP-RIBOSYLATION]. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -|- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor -
CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.
CC -|- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -|- SUBCELLULAR LOCATION: Nuclear.
CC -|- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A
CC SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -|- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES
CC OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM
CC IN LATER EMBRYOS.
CC -|- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN
CC EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.
CC -|- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.
CC -|- SIMILARITY: BELONGS TO THE PAP FAMILY.
CC -|- SIMILARITY: CONTAINS 1 BRCT DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
```



CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----

CC EMBL; D13806; BAA02964.1; -  
CC EMBL; AF051548; AAC24518.1; -  
CC EMBL; AF051548; AAC24518.1; JOINED.  
CC EMBL; AF051545; AAC24518.1; JOINED.  
CC EMBL; AF051546; AAC24518.1; JOINED.  
CC EMBL; AF051547; AAC24518.1; JOINED.  
CC EMBL; AE002935; AAF45400.1; -  
CC EMBL; AE002666; -; NOT\_ANNOTATED\_CDS.  
CC EMBL; AE002892; AAF45445.2; ALT\_SEQ.  
CC PIR; A47474; A47474.  
CC HSP; P26446; 1A26.  
CC FlyBase; FBgn0010247; Parp.  
CC InterPro; IPR001357; BRCT.  
CC InterPro; IPR001290; PARP.  
CC InterPro; IPR004102; PARP\_reg.  
CC InterPro; IPR001510; Znf-PARP.  
CC Pfam; PF00533; BRCT; 1.  
CC Pfam; PF00644; PARP; 1.  
CC Pfam; PF02877; PARP\_reg; 1.  
CC Pfam; PF00645; Znf-PARP; 2.  
CC ProDom; PD004675; Znf-PARP; 2.  
CC SMART; SM00292; BRCT; 1.  
CC PROSITE; PS00172; BRCT; 1.  
CC PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 1.  
CC PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
CC Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.  
FT DNA\_BIND 1 367  
FT DOMAIN 368 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 380 454 BRCT.  
FT DOMAIN 508 994 NAD-BINDING.  
FT ZN\_FING 19 54 PARP-TYPE.  
FT ZN\_FING 123 161 PARP-TYPE.  
FT DOMAIN 208 210 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 223 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT VARSPIC 376 564 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 994 AA; 113791 MW; ACA85A270DD29E08 CRC64;

Query Match 32.7%; Score 979.5; DB 1; Length 994;  
Best Local Similarity 38.3%; Pred. No. 1.9e-56;  
Matches 215; Conservative 110; Mismatches 199; Indels 37; Gaps 11;

Qy 12 GRARALNESRVNGNTAPEDSSPAKTRCQOE--SKMPVAGGRANKDRTEKQDES 69  
Db 463 GAIKVIN-STCISWGTDPKSRIPKWTKSLNSIYTKSPVS----- 505  
Qy 70 VKALLKGAPVDECTAKGKAHVYCEGNDVYDMLNOTLNQNNKYYLIQLLEDDAQ 129  
Db 506 -RTEKVRDGLAVDPD-SGLEIDIAHVYDNNKYSVVLGLTDIQRNKNYSKVQLLKADKK 563  
Qy 130 RNFSVWRWGRVGMGOHSLVACSGNLNKAKEIPQKFLDKTKNNWEDREKFEKVPKYD 189  
Db 564 EKYWIFRSWGRIGTNGICNSLEEDFTSESARKNFKETADTKGNEYQRDNFVKRTGRMY 623  
Qy 190 MLDYATNTQDEETKESKSLKPLKPSQLDLRVQLKICNVQAMEMMMKYNT 249  
Db 624 PIETQY-----DDDKLVKSHSFF-----TSKLEISVQNLKILIPDIISMKNKTLMEFHIDM 675  
Qy 250 KKAPLGLTVAIKAGYQSLKKIEDCIIRAGOHGRALMEACNEFYTRIPHDFTPLPIR 309  
Db 676 DKMPLGLSLAHQIQSAYRVKVEIYNVLECGSNTAKLIDATNFRYTLIPHNGVQLPLIE 735  
Qy 310 TQKLESEKIQLEALGDIEIAIKLVKTE-LQSPHPHLDQHYRNHLCALRPLDHSYEFKY 368  
Db 736 THQOIEDLRQMLDSLAETIAYSIKSEDSVSDACNPLDNHYAQIKTQLVALDKNSEEFSI 795

Qy 369 ISOYQSTHAPTSHDYTMTLLDLFEVKGDEKFAFR--EDLHNRMLLWHGSRMNVGIL 426  
Db 796 LSQVKNTHASTHKSVDLKIVDFKVSQRQGEARREKFKLHNRKLLWHGSRILNFVIGIL 855  
Qy 427 SHGLRIAPPAPITGYMFGKIYFADMSKSSANYCFASRLKNTGLLLSEVALGOCNELL 486  
Db 856 SHGLRIAPPAPITGYMFGKIYFADMSKSSANYCCTSQNSTQSNSTGLMLLSEVALGDMMBCT 915  
Qy 487 EANPKAEGLOGKHSTKGLGKMAPS-SAHFVTLNGSVPLGSPASDTGTLNPDGVTLYNNE 545  
Db 916 SAK-YINKLSNNKSCFGRTMPDPTKSYIRSDGVEIPYGETITDEHLKS---SLLYNE 971  
Qy 546 YIVYNPNQVRMYRLKVKQFNF 566  
Db 972 YIVYDVAQVNIQYLFMEFKY 992

# RESULT 12

ID PPOL\_SARPE STANDARD; PRT; 996 AA.  
AC Q11208;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
DE ribosyltransferase (Poly[ADP-ribose] synthetase).  
OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Oestroidea; Sarcophagidae; Sarcophaga.  
OX NCBI\_TaxID=7386;  
[1]  
RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=94170813; PubMed=8125121;  
RA Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,  
RA de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;  
RT "Cloning and functional expression of poly(ADP-ribose) polymerase  
RT cDNA from Sarcophaga peregrina";  
RL Eur. J. Biochem. 220:607-614(1994).  
CC -|- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY[ADP-RIBOSYL]ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -|- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor -  
CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.  
CC -|- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -|- SUBCELLULAR LOCATION: Nuclear.  
CC -|- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -|- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -|- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D16482; BAA03943.1; -  
CC HSP; P26446; 1A26.  
CC InterPro; IPR001357; BRCT.  
CC InterPro; IPR001290; PARP.  
CC InterPro; IPR004102; PARP\_reg.  
CC Pfam; PF00533; BRCT; 1.  
CC Pfam; PF00644; PARP; 1.

```
DR Pfam: PF02877; PARP_reg; 1.  
DR Pfam: PF00645; zf-PARP; 2.  
DR ProDom: PD004675; zf-PARP; 2.  
DR SMART: SM00292; BRCt; 1.  
DR PROSITE: PS0172; BRCt; 1.  
DR PROSITE: PS00347; PARP_ZN_FINGER_1; FALSE_NEG.  
DR PROSITE: PS00664; PARP_ZN_FINGER_2; 2.  
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
ADP-ribosylation; Zinc-finger; Zinc.  
FT DNA_BIND 1 369 BY SIMILARITY.  
FT DOMAIN 370 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 382 456 BRCt.  
FT DOMAIN 508 996 NAD-BINDING.  
FT ZN_FING 19 54 BY SIMILARITY.  
FT ZN_FING 126 164 BY SIMILARITY.  
FT DOMAIN 211 214 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 232 235 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
SQ SEQUENCE 996 AA; 113018 MW; 690DD36E7487298 CRC64;  
  
Query Match 32.2%; Score 964; DB 1; Length 996;  
Best Local Similarity 39.9%; Pred. No. 2e-55;  
Matches 216; Conservative 116; Mismatches 177; Indels 32; Gaps 13;  
  
QY 33 SSPAKTRRCQRESKMPVAGGKANKKRDTEKQDESVALLLKGPVDPDPECTAKVGKA 92  
Db 480 TDPATRTQESKSSKSIYTKSVPKSMT-----LKIKDGLAVDPD-SGLEDDVA 528  
  
QY 93 HVYCEGN-DVYDVLNQTNLQFNNNKYYLIQLLEDDAQRNFSVMMRGVGMKGHSLV 150  
Db 529 HVYVRNKKRYNVLGTTDIQKNKSNFYKLQLESMDKNRFFWFRSGRIGTTIGGNKLD 588  
  
QY 151 ACSGNLNAKEIFQKKFLDKTKNNWEDREKFEKVPCKYDMLQMDYATNTQDEETKES 210  
Db 589 NFS-NLVDALVQKELYLEKSGHNFENRFVAVGRMTPIDIDYA-----EDSKIDIS 641  
  
QY 211 LKSPKPESQLDLRVQELIKLICNVQAMEMMEMKYNTPKAPGLKLTVAQIKAGYQSLK 270  
Db 642 AEHDIK--SKPLSVQDIILKMFVDSMKRTMFEFLDMKPLGKLSQKIQSAVKVLT 699  
  
QY 271 KIEDCIRAGQGRALMEACNEFYTRPHDEGLTPTPLRTQKELSEKIQLEALGDIETA 330  
Db 700 EYIELQGGGTNAKFDATNRFTYLPHPNGTQSPPLDTEQVEQLQMLDSLIEICA 759  
  
QY 331 IKLVKTELOSPE-HPLDQHYRHLICALPLDHSYEFKYSYLOQTHTAPTHSDYTMILL 389  
Db 760 YSLQTEDSKADINPDKHQYQKLEPLDKNSEYILLQKYVKNTHAETHKLYLDEVV 819  
  
QY 390 DLFEVEKDGKEAFR--EDLHRLMLWHGSRMSNWWGILSHGLRIAPPAPITGYMFGKG 447  
Db 820 DIFKVARQGEARRYKPKFKLHNRRLWHGSRITNFAGILSHGLKAPPEAPVTGYMFGKG 879  
  
QY 448 IYFADMSKSNYCFASRLKNTGLILLSEVALQCOHLELANPKAGLLOGKHSTGLGK 507  
Db 880 IYFADMVSKSNYCCYSHHNSHTGLMLLSEVALGDMMECTAAKYVTK-LPNDKHSCFGR 938  
  
QY 508 MAPS-SAHFVTGLNGSVPLG-PASDTGILNPDGTYTINYEYIIVNPQVMRYLLKVVQFN 565  
Db 939 TMEPNSESIREDGVEPLGKPTINDSLKS-----SLLYNEFIYDIAQVNIQVLMRNFK 994  
  
QY 566 F 566  
Db 995 Y 995  
  
RESULT 13  
ID PPO3_HUMAN STANDARD; PRT; 533 AA.  
AC Q9Y6F1; Q9UG81;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly (ADP-ribose) polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+)
```

```
DE ribosyltransferase-3) (Poly(ADP-ribose) synthetase-3) (PADPRT-3)  
GN (hPARP-3).  
OS ADPRTL3 OR PARP3 OR ADPRT3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Fetal brain;  
RX MEDLINE=99263509; PubMed=10329013;  
RA Johansson M.;  
RT "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA  
cloning of two novel poly(ADP-ribose) polymerase homologues.";  
RL Genomics 57:442-445(1999).  
RN [2]  
RP SEQUENCE OF 75-533 FROM N.A.  
RC TISSUE=Kidney;  
RA Ansorge W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;  
Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose](N)-acceptor =  
nicotinamide + [ADP-D-ribose](N+1)-acceptor.  
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).  
CC -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in  
the kidney, skeletal muscle, liver, heart and spleen; also  
detected in pancreas, lung, placenta, brain, leukocytes, colon,  
small intestine, ovary, testis, prostate and thymus.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC  
CC EMBL: AF083068; AAD29855.1; -;  
DR EMBL: AL050034; CAB43246.1; -;  
DR HSP: P26446; I426  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR004102; PARP_reg.  
DR Pfam: PF00644; PARP; 1.  
DR Pfam: PF02877; PARP_reg; 1.  
KW Transferase; Glycosyltransferase; NAD; Nuclear protein;  
ADP-ribosylation.  
FT DOMAIN 14 20 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
FT CONFLICT 80 80 K -> N (IN REF. 2).  
FT CONFLICT 171 171 A -> G (IN REF. 2).  
FT CONFLICT 411 411 K -> E (IN REF. 2).  
SQ SEQUENCE 533 AA; 60117 MW; 7C0AB89E64D1B9FD CRC64;  
  
Query Match 24.0%; Score 720; DB 1; Length 533;  
Best Local Similarity 34.8%; Pred. No. 7.3e-40;  
Matches 193; Conservative 98; Mismatches 188; Indels 76; Gaps 20;  
  
QY 46 ESKKMPVAGGKANKKRDTE---KQDESVALLLKGR-APVDPECTAKVGK-AHVYCEGND 100  
Db 15 EKKGQAG-----REDFPRSTAALKAIPAERIIIRDVPTCLSSNPGTQVY----E 64  
  
QY 101 VYDVLNQTNLQFNNNKYYLIQLLEDDAQRNFSVMMRGVGMKGHSLVACSGNLNKA 160  
Db 65 DYNCTLNQTNINNKKFYIQLLQ-DSNRFFTCMNRGRVGEVGQ-SKINHPTRELDAR 122  
  
QY 161 EYFQKFLDKTKNNWEDREKFEKVPCKYDMLQMDYATNTQDEETKESLKLPLPESQ 220  
Db 123 KDFEKKFREKTKNNWAEHDHFVSHPGKYTLIEV-----QADEAQAQAVKVDAPVPTVK 178  
  
QY 221 -----LDLRVQELIKLICNVQAMEMMEMKYNTPKAPGLKLTVAQIKAGYQSLK 274  
Db 179 RVQPCSLDPATQKLTNIFSKEMFKYNTMALMDLVKMKPLGKLSKQIARGFEALEEE 238
```

QY 275 CIRA-GQHRALMEACNEFYTRIPIHDFGLRTPPLIRTOQKELSEKIQLEALGDIEI--AI 331  
 Db 239 ALKQPTDQSGSLEELSSHYFVPHNFHSGPPPLNSPELLQAKDMLLADIELAQA 298  
 QY 332 KLKTELQSP--HPDQHYRNHLCALRPDLHSEYEFKVLISQYLOSTHATHSDYTWL 388  
 Db 299 QAVSEQKTEVEVPHPLDRDQOLKQQLDLSGAPYKVIQTYLEQFGS---NHRCPTL 355  
 QY 389 LDLEVEKDGKEAFR--EDLHNRMLWHGSRMSNWVGLSHGLRIAPPEAPITGYMEFK 446  
 Db 356 QHINKVNGEEDREFQAHKSLGNKLLWHGTMVAVAILTSGLRIMPH-----SGGRVGK 411  
 QY 447 GIYFADMSKSNKYCFASR--LKNLTGGLLLSEVALGQCNELEANPKRAEGLLQKHSTKG 504  
 Db 412 GIYFASENSKSAGYVIGMKCAHGVYMFGLGEVALGREHINTDNPDL----- 459  
 QY 505 LGKMAPSSAHFTVLTNGSVPLGPASDT-----GILNPDGY-----TLNNEYI 547  
 Db 460 --KPPPGDFSVIARGHTPEP-DPTQDTELDGQGVVPOGVPCPEFSSTFSQSEYL 516  
 QY 548 VYNPNQVRMYLLKV 562  
 Db 517 IYQESQCLRYLLEV 531  
 RESULT 14  
 YQNA\_CAEEL STANDARD; PRT; 538 AA.  
 AC Q09525;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-FEB-1996 (Rel. 33, Last annotation update)  
 DE Hypothetical 61.3 kDa protein E02H1.4 in chromosome II.  
 GN E02H1.4.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BRISTOL N2;  
 RA Smith A.;  
 RL Submitted (JAN-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: SOME, TO THE MIDDLE PART OF NAD(+) ADP-  
 CC RIBOSYLTRANSFERASE (EC 2.4.2.30).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; 247075; CAA87379.1; -  
 DR HSSP; P26446; 1A26.  
 DR WormPep; E02H1.4; CE01539.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 538 AA; 61268 MW; 3144E25465FC7341 CRC64;

Query Match 17.5%; Score 526; DB 1; Length 538;  
 Best Local Similarity 29.1%; Pred. No. 3.5e-27;  
 Matches 158; Conservative 103; Mismatches 182; Indels 100; Gaps 22;  
 QY 102 YDVMNLQTNLFQNNKYYLIQLLEDQARNFVWNRGRVCKMGQHSILVACSGNINRAKE 161  
 Db 12 YKVHLCKNTIAQNNKFDYMLLEDGG--DFIVKLINGRIGRYGTQLKDFD-DLDRACK 68

QY 162 IFQKKFLDKTKNNWEDREKFEKVPKQYMLQMDYATNT-QDEEETKKEE----- 209  
 Db 69 FFESEKYEKTHLWHERDD-EPVFNKYAVAVEL--ATNARQTEKEVKEPEPEPKVDEKN 125  
 QY 210 -----SLKSPLPAPESQDLRLVQELIKLIGNVQAMMEMMYNTK-KAPL 254  
 Db 126 TRGRRKRGIVKEKEIKKEEPEVEVEVEVEVEVEVEVEVEVEVEVEVEVEVEVEVEVEVE 185  
 QY 255 GKLTVAQIKAGYQSLKKTEDCI-----RAGQHGR-----ALMEACNEFYTR 295  
 Db 186 DCLSLAQLTQYEILSKIEESIGKKSARRSTRGRPRVADRVLAVKSDGSLHDINKYISL 245  
 QY 296 IPHDFGLRTPPLIRTOQKELSEKIQLEAL-GDIEIAIKLVKTELQSPHPLDQH-YRNLH 353  
 Db 246 IPHSGFCVPPKIDSHAKIQAEERELLDALKGSIASLEALDKLKKTKTASSKDIYQRLYERLP 305  
 QY 354 CALRPLDHSEYEFKVISQYL--QSTHAPTHSDYTWLTLDFEVEKDE----- 399  
 Db 306 CHLEFPVSE-----IAGKIGDCLAMRGPTHG-YKLSLIDAFELKDPNEIPTAPVEVQE 358  
 QY 400 --KEAFREDLH-----NRLLWHGSRMSNWVGLSHGLRIAPPEAPITGYMEFKG 447  
 Db 359 VPKRGKRSKTYAAPTVPPTTKRLWHGTRVTVNVSILMGLQF--PVGRCGLMFGNG 416  
 QY 448 IYFADMSKSNKYCFASRLKNTGLLLSEVALGQCNELEANPKRAEGLLQ--GRHSTKGL 505  
 Db 417 VYFANVPTKSANYCCPEASKRV-FMLLCEVETANPLVLYESEIDADEKMEKAKKTSVAA 475  
 QY 506 GKMAPSSAHFTVLTNGSVPLGPASDTGILNPDGYT-LYNNEYIYVNPQVRMYLLKVKQF 564  
 Db 476 GKHTPRDT--VEINGI-----PAFKSNLEETIEETRELLYDEYVNFENKHEFKYVVEVKV 528  
 QY 565 NFL 567  
 Db 529 DRL 531  
 RESULT 15  
 PPOL\_ONCMA STANDARD; PRT; 135 AA.  
 AC Q08824;  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly (ADP-ribose) polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
 DE ribosyltransferase (Poly(ADP-ribose) synthetase) (Fragment).  
 GN ADPRT.  
 OS Oncorhynchus masou (Cherry salmon) (Masu salmon).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;  
 OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.  
 OX NCBI\_TaxID=8020;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93277538; PubMed=8503897;  
 RA Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,  
 RA Miwa M.;  
 RT "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)  
 RT polymerase from Xenopus laevis and cherry salmon using heterologous  
 RT oligonucleotide consensus sequences.";  
 RL Biochem. Biophys. Res. Commun. 193:119-125(1993).  
 CC -1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribose}(N)-acceptor =  
 CC nicotinamide + {ADP-D-ribose}(N+1)-acceptor  
 CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,  
 CC AND BRAIN. LOW IN LIVER.

-I- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN LENGTH OF 20-30 UNITS.

-I- SIMILARITY: BELONGS TO THE PARP FAMILY.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; D13809; BAA02965.1; -.  
PIR; PNO494; PNO494.  
HSSP; P26446; 1A26.  
InterPro; IPR001290; PARP.  
InterPro; IPR001510; Znf-PARP.  
Pfam; PF006644; PARP; 1.  
PROSITE; PS00347; PARP\_ZN\_FINGER\_1; PARTIAL.  
PROSITE; PS50084; PARP\_ZN\_FINGER\_2; PARTIAL.  
Transferase; Glycosyltransferase; NAD;  
Adp-ribosylation; Zinc-finger; Zinc.  
NON\_TER 1 1  
FT DOMAIN <1 >135 NAD-BINDING.  
FT ACT\_SITE 135 135 BY SIMILARITY.  
FT NON\_TER 135 135  
SQ SEQUENCE 135 AA; 15411 MW; A60E8E98890E42DC CRC64;

Search completed: August 29, 2002, 08:02:11  
Job time: 288 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:55:28 ; Search time 117.38 Seconds  
(without alignments)  
840.067 Million cell updates/sec

Title: US-09-701-586b-2

Perfect score: 2998

Sequence: 1 MAARRRRSTGGRRALNES.....PNOVRMYLLKVFNLQLW 570

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1176	39.2	635	10 O81294	O81294 arabidopsis
2	1149.5	38.3	653	10 O50017	O50017 zea mays (m
3	1022	34.1	607	13 QPS82	QPS82 gallus gall
4	1018	34.0	1014	11 Q921K2	Q921K2 mus musculus
5	988	33.0	612	5 Q9TX06	Q9TX06 dictyosteli
6	981.5	32.7	607	13 Q9PS81	Q9PS81 xenopus. na
7	972.5	32.4	593	5 Q9TX05	Q9TX05 drosophila
8	966	32.2	983	10 Q92P54	Q92P54 arabidopsis
9	966	32.2	1009	10 Q9SJW4	Q9SJW4 arabidopsis
10	961.5	32.1	969	10 O24570	O24570 zea mays (m
11	961.5	32.1	980	10 Q9ZSV1	Q9ZSV1 zea mays (m
12	777.5	25.9	945	5 Q9N4H4	Q9N4H4 caenorhabdi
13	725	24.2	533	4 Q96CG2	Q96CG2 homo sapien
14	696.5	23.2	528	11 Q91YR6	Q91YR6 mus musculu
15	685.5	22.9	727	5 Q9X0A5	Q9X0A5 caenorhabdi
16	408.5	13.6	815	10 Q9FK91	Q9FK91 arabidopsis

17 396.5 13.2 815 10 Q9SWB4  
18 353.5 11.8 2276 5 Q9TXQ1  
19 141.5 4.7 181 12 O55721  
20 135 4.5 840 6 Q95JR0  
21 135 4.5 1327 4 Q95271  
22 134.5 4.5 1170 16 Q9X0R4  
23 134.5 4.5 1933 5 Q95XK2  
24 133.5 4.5 1065 2 Q9AHK9  
25 128.5 4.3 1553 5 Q27421  
26 128 4.3 1181 5 Q9XZ37  
27 128 4.3 1181 5 Q9VBP3  
28 127.5 4.3 479 2 Q9LAX2  
29 125.5 4.2 481 2 Q9LAX5  
30 125 4.2 962 2 Q49546  
31 125 4.2 1051 2 Q49524  
32 125 4.2 1365 2 Q49525  
33 123 4.1 1173 2 Q9AHK6  
34 122.5 4.1 957 2 Q9AHL1  
35 122.5 4.1 3130 5 Q9BK46  
36 122 4.1 1065 2 Q9AHK8  
37 121 4.0 533 6 Q95JY2  
38 121 4.0 753 4 Q09471  
39 121 4.0 1166 4 Q9H2K2  
40 121 4.0 1206 4 Q9NFC1  
41 121 4.0 1265 4 Q9HAS4  
42 121 4.0 1820 4 Q9GQ89  
43 120.5 4.0 469 4 O00467  
44 120.5 4.0 1270 5 Q20749  
45 120.5 4.0 1931 13 Q910C5

#### ALIGNMENTS

RESULT 1  
O81294 PRELIMINARY; PRT; 635 AA.  
ID O81294;  
AC O81294;  
DC 01-NOV-1998 (TrEMBLrel. 08, Created)  
DT 01-NOV-1998 (TrEMBLrel. 08, Last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DEF T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).  
OS Arabidopsis thaliana (Mouse-ear cross).  
GN T14P8.19 OR AT4G02390.  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA WASHU;  
RT "The A. thaliana Genome Sequencing Project.";  
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. COLUMBIA;  
RA Kalicki J., Elliott G., Cloud J.;  
RT "The sequence of A. thaliana T14P8.";  
RN [3]  
RP Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
RC STRAIN=CV. COLUMBIA;  
RA Waterston R.;  
RT Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC Lamar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X.;  
RT Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;

Q9swb4 glycine max  
Q9txq1 caenorhabdi  
O55721 chilo iride  
Q95jr0 macaca fasc  
Q95271 homo sapien  
Q9x0r4 thermotoga  
Q95xk2 caenorhabdi  
Q9ahk9 borrelia bu  
Q27421 drosophila  
Q9x237 drosophila  
Q9vbp3 drosophila  
Q9lax2 streptococc  
Q9lax5 streptococc  
Q49546 mycoplasma  
Q49524 mycoplasma  
Q49525 mycoplasma  
Q9ahk6 borrelia bu  
Q9ahl1 borrelia bu  
Q9bhk8 borrelia bu  
Q95jy2 macaca fasc  
Q09471 homo sapien  
Q9h2k2 homo sapien  
Q9ntc1 homo sapien  
Q9has4 homo sapien  
Q96q89 homo sapien  
O00467 homo sapien  
Q20749 caenorhabdi  
Q910c5 gallus gall

RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF069298; AAC19283.1; -  
DR EMBL: AL161494; CAB80732.1; -  
DR HSSP: P26446; 1A26  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR004102; PARP\_reg.  
DR InterPro: IPR003034; SAP.  
DR Pfam: PF00644; PARP; 1.  
DR Pfam: PF02877; PARP\_reg; 1.  
DR Pfam: PF02037; SAP; 2.  
DR SMART: SM00513; SAP; 2.  
DR Transferrase.  
SQ SEQUENCE 635 AA; 72017 MW; E3F1CBE4D367A377 CRC64;

Query Match 39.2%; Score 1176; DB 10; Length 635;  
Best Local Similarity 43.9%; Pred. No. 5.7e-74;  
Matches 254; Conservative 98; Mismatches 187; Indels 40; Gaps 11;

QY 2 AARRRSTGGGR---ARALNESKRVNNGNTAPEDSSPAKTRRCQROESKMPVAGGKA 57  
DB 81 AAKRGDITGTTKDLRLCNDAVN---APVKS-----NDEA 118  
QY 58 NKDRT---EDKQDESVALLLKGAIPVDPCTAKV-GKAHYVCEGNDVYDVMNLQTNLF 113  
DB 119 EDDNNGFEKEKEEIVTATKGAVALDQWIPDEIKSQYHVLQRGDDVYDAILNQTNRD 178  
QY 114 NNKYYLIQLLEDDAQRNSVMRWGRVGMQSHSLVACSGNLNKAKEIFQKFLDKTKN 173  
DB 179 NNKFFVLQVLESDDSKTKTYVYTRWGRVGVGQSKLDGPDYDSMDRAIEFTNKNDKTKN 238  
QY 174 NWEDREKFEKPGKYDMLQMDATNTQDEETKKEESLSPKPE-SQLDLRVQELIKLI 232  
DB 239 YNSDRKEFTPHKPSYTWLEMDYKKEENDSPVNDIPSSSEVKGPEQSKLDTRVAFISLI 298  
QY 233 CNVQAEEMMMKMYTKKAPLGKLTVAQIKAGYQSLKKIEDICIRAGQHGRALMEACNEF 292  
DB 299 CNVSMQAHMMEIGYNANKLPLGKISKSTISKGYEVLRKRISEVIDRYDRTR-LEELS 357  
QY 293 YTRIPHDGLR--TPPLITQKELSEKIQLEALGDIETAIKLVKTELQSPHPLDQHYR 350  
DB 358 YTVIPHDGFKMSQSVVIDTPQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQ 417  
QY 351 NLHCLARPLDHESEYFKVISOYLSQTHAPTHSDYTWLTLDFEVEKDGKEAPRE--DLH 408  
DB 418 QNLGCLTPVNDSEESVMYANTHENTHATGSHYTVETIAQLFRASRAVDAPQFSSSK 477  
QY 409 NRMLLWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKGIYFADMSKSNKYCFASRLKN 468  
DB 478 NRMLLWHGSRMTNWAGILSGLRIAPPEAPITGYMFGKGIYFADMSKSNKYCFANTGAN 537  
QY 469 TGLLLSEVALGOCNELLKANPKAELLOGKHSKGLGKMAPSSAHFVTL-NGSTVPLGP 527  
DB 538 DGVLLCEVALGDMNELLSDYNADNLPGLKSLTKGVTAPNPSAOTLEDGVVYVPLGK 597  
QY 528 ASDTGILNPDGTYTLNNEYIVNPNQVRMYLLKVQNF 566  
DB 598 PVERSCSKG---MLLYNEIVNVEQIKMYIVQVKNFY 633

RESULT 2  
ID O50017 PRELIMINARY; PRT; 653 AA.  
AC O50017;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE POLY(ADP-RIBOSE) POLYMERASE.  
GN PARP.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.

OX NCBI\_TaxID=4577;

RN [1]  
RP SEQUENCE FROM N.A.  
RA Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,  
RT O'Farrell M., Van Montagu M., Inze D., Kushnir S.;  
RL "Higher plants possess two poly(ADP-ribose) polymerases.";  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ222588; CAA10888.1; -  
DR HSSP: P26446; 1A26.  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR004102; PARP\_reg.  
DR InterPro: IPR003034; SAP.  
DR Pfam: PF00644; PARP; 1.  
DR Pfam: PF02877; PARP\_reg; 1.  
DR Pfam: PF02037; SAP; 2.  
DR SMART: SM00513; SAP; 2.  
SQ SEQUENCE 653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

Query Match 38.3%; Score 1149.5; DB 10; Length 653;  
Best Local Similarity 44.0%; Pred. No. 4.2e-72;  
Matches 255; Conservative 89; Mismatches 191; Indels 45; Gaps 13;

QY 1 MAARRRSTGGGRARALNESKRVNNGNTAPEDSSPAKTRRCQROESKMPVAGGKAND 60  
DB 102 LAKARGVAANGCKDVQIRLLSATAGPAAVADGGPL-----GAKEVIKGG---DE 148  
QY 61 RTEDQDESVALLLKGAIPVDPCTAKV-GKAHYVCEGNDVYDVMNLQTNLFNNKYY 119  
DB 149 EYEVKKEKMTVA-TTKGAVALDQHPDHKVNHYHVLQVGDIEATLQNTNGDNNKFEY 207  
QY 120 LTOLLEDDAQRNSVMRWGRVGMQSHSLVACSGNLNKAKEIFQKFLDKTKNWEDERE 179  
DB 208 IIQVLESADGSGFMYNWRGVGVGQDKLHGPSRTRDQAIYEFEGKFHNKHNHSDRK 267  
QY 180 KFEKPGKYDMLQMDATNTQDEETKKE---ESLSKPLKPESQDLRLVQELIKLICNVQ 236  
DB 268 NFKCYAKKYTWLEMDY-----ETKEIEKGSITDQIK-ETKLETRIAQFISLICNIS 319  
QY 237 AMEEMMMKMYTKKAPLGKLTVAQIKAGYQSLKKIEDICIRAGQHGRALMEACNEFYTRI 296  
DB 320 MNKQRMVEIGYNAELPLGKRLKATILKGYHVLKRTISDVISKADR-RHLEQTLGEFTVI 378  
QY 297 PHDFGLRTPP--LIRTKELSEKIQLEALGDIETAIKLVKTELQSPHPLDQHYRNLC 354  
DB 379 PHDFGFRKREFIIDTPQKLKAKLEWVEALGEIETATKLLDESDDDPLARYKQLHC 438  
QY 355 ALRPLDHESEYFKVISOYLSQTHAPTHSDYTWLTLDFEVEKDGKEAPRE--DLHNRML 412  
DB 439 DFTPLEADSDEYSMTKSYLRNTHGKTHSGYTVDIQVIFKVSRRHGETERFQKEASTNRML 498  
QY 413 LWHGSRMSNWGILSHGLRIAPPEAPITGYMFGKGIYFADMSKSNKYCFASRLKNTGLL 472  
DB 499 LWHGSRLSNWAGILSGLRIAPPEAPITGYMFGKGIYFADMSKSNKYCFASRLKNTGLL 558  
QY 473 LLSEVALGOCNELLKANPKAELLOGKHSKGLGKMAPSSAHF-VTLNGSTVPLG----- 526  
DB 559 LLCEVALGDMNELLNADYDANNLPKGLKSKGVGQTAPNHWESKVDGCVVYVPLGPKQE 618  
QY 527 PASDTGILNPDGTYTLNNEYIVNPNQVRMYLLKVQNF 566  
DB 619 PSKRGGLL-----YNEYIVNVDQIRMVYLVHNFNF 650

RESULT 3  
ID Q9PS82 PRELIMINARY; PRT; 607 AA.  
AC Q9PS82;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE NAD:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.  
OS Gallus gallus (Chicken).



RA	Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA	Kofler B., Schweiger M., Wagner E.F.;
RT	"On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT	riboseyl) transferase (ADPRT): ADPRT from Dictyostellium discoideum and
RT	inactivation of the ADPRT gene in the mouse.";
RL	Biochimie 77:444-449(1995).
DR	HSP: P26446; 1A26.
DR	InterPro; IPR001290; PARP.
DR	InterPro; IPR004102; PARP_reg.
DR	Pfam; PF00644; PARP; 1.
DR	Pfam; PF02877; PARP_reg; 1.
SQ	SEQUENCE 612 AA; 69241 MW; CB340F7A88FE2364 CRC64;

Query Match	33.0%;	Score 988;	DB 5;	Length 612;
Best Local Similarity	39.6%;	Pred. No. 7.4k-61;		
Matches 230;	Conservative 100;	Mismatches 207;	Indels 44;	Gaps 16;
Qy	1	MAARRRSTGGRRALNFKSRVNGNTAPEDSSPAKTRRCORQESKKMPVAGGKANKD	60	
Db	57	VAKPTRSSKKAPAKSEKGEKXNKLTT-----RPTKTKTNTSAETKTAASDLDDSSS	111	
Qy	61	RTEDKQDE-SYKALLLKAKAPVDPECTAKVGAHAYCEGNDVYVMLNQTNLQFNNNKYY	119	
Db	112	ESEDEKNQISVK---IKGRAANDPHFPD--SRXHVYENGKDVYDATLNAQTEISQNNKKY	166	
Qy	120	LIQLEDDAQRNFVWMRGVRGKMGQHSILVAC-SGNLNKAKEIFQKFKLOKTKNNWEDR	178	
Db	167	IIQLEADGGSYSVWNWNRGELGKQSRDKFCGGLNGQAISLFCSFYFKTKNTFTFR	226	
Qy	179	EKFKEVPCKYDMLQWDYATNTQDE---EETKKEESLKSPL--KPESOLDLRVQELIKLIC	233	
Db	227	ANFKKVGAKYDMIELDYSTDSPKNGASTATTTTKKVVEHKECSLDERVQELVKLIF	286	
Qy	234	NVQAMEEEMEMKYNTKAPIGKGLTVAQIKAGYOSLKKIEDICIRAGOHGRALMEACNEFY	293	
Db	287	DVKAMERTMTAKYDLKKMPGLKLSKNOITKGYLVLKQIED-VMGKSGESLSTLSRRFY	345	
Qy	294	TRPHDFGRPTPLIRTOKESELKOLLEALGDIEATIKLVKTEQLQSPHEPLDQHYNLH	353	
Db	346	TIIPHAFGMSVPPVINTNOMLIEKNMLQNLOADIEATNIIKDESDESNILELHYAKLK	405	
Qy	354	CALRPLDHESEYFKVISOYLOSTHAPTHSDYTMTLDDLFEVEKDGEKBAFREDLH--NRM	411	
Db	406	TDIQPLDENSECYKNILLYVKNY---QGGKKPTIVNIFKIDRQGEADRYKTKKHLGNRK	462	
Qy	412	LLWHGRMSNVWGIILSHGLRTAPPEAPITGYMFGKGIYFADMSKSNACYFASRLKNTGL	471	
Db	463	LLWHGSRLTNVASHIISQGLRLTAPPEAPVSGYRFGKGYFADCMLSANYS-----RTVGF	517	
Qy	472	L-----LLSEVALGQCNELLEANPKAEGLLQGGKTKGLGKNAPS-SAHFVTLNGSTVPL	525	
Db	518	LIDFCMLLGDVALGKTADLAR-DIYMEKQPQNSHSTWALGTVEPDPKVFQDTGECITPY	576	
Qy	526	G----PASDTGILNDPGTYTLNTNVEIVYNPNQVRMYLLKVQ	563	
Db	577	GQMTPSQHKGY-----SCYEHQTVVYDVAQVHLKYLQLR	611	

RESULT	6
Q9PS81	
ID	Q9PS81 PRELIMINARY; PRT; 607 AA.
AC	Q9PS81;
DT	01-MAY-2000 (TREMBLrel. 13, Created)
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT	01-DRC-2001 (TREMBLrel. 19, Last annotation update)
DE	NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRF.
OS	Xenopus.
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC	Xenopodinae.
OX	NCBI_TaxID=8353;
RN	[1]

SEQUENCE FROM N.A.  
 RX MEDLINE=96007847; Pubmed=7578427;  
 RA Auer B., Flick K., Wang Z.Q., Haldacher D., Jager S., Berghammer H.,  
 RA Kofler B., Schweiger M., Wagner E.F.;  
 RT "On the biological role of the nuclear polymerizing NAD<sup>+</sup>: protein(ADP-  
 ET ribosyl) transferase (ADPRT): ADPRT from Dictyostellum discoideum and  
 RT inactivation of the ADPRT gene in the mouse."  
 RL Biochimie 77:444-449(1995).  
 DR HSP; P26446; 1A26.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR PROSITE; PS01172; BRCT; 1.  
 SQ SEQUENCE 607 AA; 67496 MW; 54CDBBBE22079886 CRC64;

Query Match	32.7%	Score	981.5	DB	13	Length	607
Best Local Similarity	40.9%	Pred.	No. 2.le-60				
Matches	216	Conservative	101	Mismatches	80	Indels	31
Gaps	13						
Qy	51	PVAG---GKANKRTEKQDSV	KALLK	KAPVDP	ECTAKV	GAHVC	YCEGNDV
Db	93	PVAGSGKVKKEGSKNSK	SEKMK-L	TVKGAA	ADP	DOSELE	-DSCHVLE
Qy	108	QTNLFQNNKYYLQLLE	DDAQR	NFSVM	WRGVK	-MGQHS	L
Db	151	LVDITRGTNSYYKL	LIETH	DRSRY	WFRSG	RVGT	YIGSKK
Qy	167	FLDKTKNWEDEK	EPVGY	KDMLQ	MDY	ANTODE	ETKKE
Db	210	YODKTGNWHS-	PNFTK	YPK	FY	PLEIDY	G---Q
Qy	227	ELIKLINVQAMEEM	EMKNT	KAPL	GLT	VAQI	RAGYO
Db	263	ELIKLIFVESMK	KAMVEF	IDLQ	MP	L	GLUK
Qy	287	EACNEFTRIPHD	FLGRT	PLIR	TQK	ELSE	LIQLE
Db	323	DLSNQFYTLIPHD	FGMKP	PLNN	LEYI	QAKV	QMLN
Qy	346	DOHYNLHUALR	PLDHS	EPKV	TSYQL	QSTH	APATH
Db	383	DVYKEIKTD	KVYAK	DSESR	ICD	VKN	THAD
Qy	405	-EDLHNRLMLH	WGSRM	NWGL	SIL	SHGL	RIAP
Db	443	FKQLHNROLMLH	WGSRT	NF	AGIL	SOGL	RIAP
Qy	463	SRKNTGLLIL	SEVAL	GOCN	ELL	SNP	KAEG
Db	502	MPGSPIGLIL	GEVAL	GNMH	-LKA	ASQIT	KL
Qy	523	SRKNTGLLIL	SEVAL	GOCN	ELL	SNP	KAEG
Db	561	MPGSPIGLIL	GEVAL	GNMH	-LKA	ASQIT	KL
Qy	566	PLGPA-----	SDT	GIL	NP	DG	VT
Db	601	PLGKGT	SANT	SDT	SILL	-----	Y

RESULT	7	
Q9TX05		
ID	Q9TX05	PRELIMINARY; PRT; 593 AA.
AC	Q9TX05;	
DT	01-MAY-2000	(TREMBLrel. 13, Created)
DT	01-MAY-2000	(TREMBLrel. 13, Last sequence update)
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)
DE	NAD+;PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.	
OS	Drosophila sp. (Fruit fly).	
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;	
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC	Ephydroidea; Drosophilidae; Drosophila.	
OX	NCBI TaxID=7242;	







```

Qy 303 RTPPLIRTKOKELSEKIOLEALGDIETAIKLVKTELQSPHPLDQHYRNLCALRPLDHE 362
Db 707 --PHIIRDEDDLMIAKAKMLEALQDIETASKIVGFDSDSE-SLDKYMKLHCDITPLAHD 763
Qy 363 SYEFKVISQYLOSTHAPSHSDYTWTLDLDFEVEKDGE---KEAFREDLHNRMLLWHGSRM 419
Db 764 SEDYKLIIEQYLLNTHAPTHKDWLSLEEVFSLDRDGLNLYSRVKNLHNRMLLWHGSRM 823
Qy 420 SNWYGILSHGLRIAPPEAPITGYMGKGIYFADMSKSNYCFASRLKNTGLLLSEVAL 479
Db 824 TNFVGILSQGLRIAPPEAPVGYMGKGIYFADLVSKSAQCYVDNRNPNVGLMLLSEVAL 883
Qy 480 GOCNELLEANPKAGLLQKHKSTGLGKMAPSSAHFVTLNGS--TVPLGPASDTGLINPDG 538
Db 884 GDMYELKAT-SMDKPPRGKHKSTGLGKTVPLESEFVKWRDDVVVPCGKPPSSIRSSE- 941
Qy 539 YTLNNEYIVYNPNQVRMYLLKVOFN 565
Db 942 --LMYNEYIVYNTSQVKMQFLKVRFH 966

RESULT 11
Q9ZSV1 ID Q9ZSV1 PRELIMINARY; PRT; 980 AA.
AC Q9ZSV1;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DE 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE POLY(ADP)-RIBOSE POLYMERASE (EC 2.4.2.30).
GN PARP1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP MEDLINE=99026291; PubMed=9808734;
RA Mahajan P.B., Zuo Z.;
RT "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL Plant Physiol. 118:895-905(1998).
DR EMBL; AF093627; AAC79704.1; -.
DR HSSP; A26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR003034; SAP.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; Zf-PARP; 2.
DR ProDom; PD004675; Znf-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS0172; BRCT; 1.
DR PROSITE; PS00064; PARP_ZN_FINGER_2; 2.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;

Query Match 32.18; Score 961.5; DB 10; Length 980;
Best Local Similarity 39.38; Pred. No. 9.9e-59;
Matches 223; Conservative 104; Mismatches 179; Indels 61; Gaps 18;

Qy 32 DSSPAKTRRCQROESKMPVAGG-----KANK-----DRTEKDQDESVKALLL 75
Db 439 ENAEVRKARL-----KIPIVREGYCECVKKNKMLPFDLKYLENALESKSGSTVTVKV 492
Qy 76 KGAPVDPCTAKVGKAHVYCEGNDYDMLNQTNFNKNNKYYLIQLEDDAQRNFSW 135
Db 493 KGRSAVH-ESSGLQDTAHILEDGKSIYNTALNMSDALGVNSYYVLQIEQDDGSECYVF 551

```

```

Qy 136 MRGRGV--KMGQSHLVACSGNLKAKEIFOKKFLDKTKNNWEDRE---KFEKVPKGYDM 190
Db 552 RKWGRVSEKIGGQKLEMS--KTEAIKEFKRFLFLEKTSWEAWECKTNFKQGRGYP 609
Qy 191 LQMDYATNQDEETKKEESLKSPLKPSQDLDRVQELIKLICNVQAEEMMEMKMYNTK 250
Db 610 LDVDYGVKAPKPKRDIS--NKSSLAPQ-----LLELMKMLFNVTYRAAMEFEINMS 661
Qy 251 KAPLGKLTVAQIKAGYQSLKKIEDCIR-AGQHGRALME-----ACNEFYTRIP--HDFGL 302
Db 662 EMPGLKLSKENIEKGFEALTEIQNLKDTADQALAVRESLIVAASNRFTLLPSIH---- 717
Qy 303 RTPPLIRTKOKELSEKIOLEALGDIETAIKLVKTELQSPHPLDQHYRNLCALRPLDHE 362
Db 718 --PHIIRDEDDLMIAKAKMLEALQDIETASKIVGFDSDSE-SLDKYMKLHCDITPLAHD 774
Qy 363 SYEFKVISQYLOSTHAPSHSDYTWTLDLDFEVEKDGE---KEAFREDLHNRMLLWHGSRM 419
Db 775 SEDYKLIIEQYLLNTHAPTHKDWLSLEEVFSLDRDGLNLYSRVKNLHNRMLLWHGSRM 834
Qy 420 SNWYGILSHGLRIAPPEAPITGYMGKGIYFADMSKSNYCFASRLKNTGLLLSEVAL 479
Db 835 TNFVGILSQGLRIAPPEAPVGYMGKGIYFADLVSKSAQCYVDNRNPNVGLMLLSEVAL 894
Qy 480 GOCNELLEANPKAGLLQKHKSTGLGKMAPSSAHFVTLNGS--TVPLGPASDTGLINPDG 538
Db 895 GDMYELKAT-SMDKPPRGKHKSTGLGKTVPLESEFVKWRDDVVVPCGKPPSSIRSSE- 952
Qy 539 YTLNNEYIVYNPNQVRMYLLKVOFN 565
Db 953 --LMYNEYIVYNTSQVKMQFLKVRFH 977

RESULT 12
Q9N4H4 ID Q9N4H4 PRELIMINARY; PRT; 945 AA.
AC Q9N4H4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 108.0 KDA PROTEIN.
GN Y71F9AL.18.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=99069613; PubMed=9851916;
RA None;
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Bradshaw-Cordum H., Scott K., Graves T.;
RT "The sequence of C. elegans cosmid Y71F9AL.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RT "Direct Submission.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AC024200; AAE36011.1; -.
HSSP; P26446; 1A26
InterPro; IPR001290; PARP.
InterPro; IPR001510; Znf-PARP.
Pfam; PF00644; PARP; 1.
Pfam; PF00645; Zf-PARP; 1.
ProDom; PD004675; Znf-PARP; 1.

```

DR PROSITE; PS50064; PARG\_2N\_FINGER\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 945 AA; 108006 MW; 1D0A62C954BC6AD9 CRC64;

Query Match 25.9%; Score 777.5; DB 5; Length 945;  
Best Local Similarity 34.7%; Pred. No. 6.8e-46;  
Matches 196; Conservative 103; Mismatches 213; Indels 53; Gaps 17;  
QY 21 KRYVNGTAPEDSSPAKTRRCORQESKMPVAGGRANKDRTEKQDESVALLLKGP 80  
DB 409 KRKSGGTRGEQFIYAAEA-----FDSTNNVPIKVGDTSTNTH-----IIRKGTV 454  
QY 81 VDPECTAKVKAHYICBGND--VYDMLNQTNLQFNNNKYLIQLLEDQAQRNFVSMRW 138  
DB 455 VDAK-FALADRCHVFRNEIDGSLYQATLSFTDLTQNKSNYKIQLLKQDQRENYVFRSW 513  
QY 139 GRVG-KMGQHSVLVACSNLKAKEIFQKFLDKTKNWNEDREKFEKVPKGYDMLQMDYAT 197  
DB 514 GRVGTGVGGNHESYS-NSNEALIKFDQVPHEKTKNDWIYRKHKMPGMFSYVETDYSE 572  
QY 198 NTO--DEEETKKEESLKSPLKPESQDLRLVQELIKLCINVOAEMEMMKYNTKKAPLG 255  
DB 573 FAQITTEITPGSKTL-----LPKSVKEVMSIFDVENMKSALKSEMDVNKMPLG 623  
QY 256 KLTVAQIKAGYQSLKKIED-CIRAGQHGRLMEACNEFYTRIPHDGFLRTPLPRTQKEL 314  
DB 624 RLSHNQINLAFVNDISDLLKLPIDASRLDFSNKFYTIIPHNFGMRVPEPIDSFHKI 683  
QY 315 SEKIQLLEALGDIEIAIKLV-----KTELQSPHPLDQHYRNLCALRPLDHESEYFK 367  
DB 684 KERNNNLALDLKFAIDQSGGDVPASTSLGI--DPVDINYOKLKCIMPLEQQGCDWN 741  
QY 368 VISQYLOSTHAPTHSDYTMILLDLFEVEKDGKEAPREDLHNRMLLWHGSRMNWVGILS 427  
DB 742 MIHQYLNKTHGATH-DLKVELIDILKLNDRNNESSKFKRHICGNRRLWHGSGKNFAGILG 800  
QY 428 HGLRIAPPEAITGYMGKGYFADMSKKSANYCFASRLKNTGLLLSEVALGQCCHELLE 487  
DB 801 QGLRIAPPEAPVSGYMGKGYFADMSKKSFFYCRAN-AKEEAYLLLCVVALGNVQQLMA 859  
QY 488 A-NPKAEGLLQGHSHKGLCKMAPSS-AHFVTLNGSTVPLGPASDTGILNPDG-----YT 540  
DB 860 SKNVSRQTLFAGQSVQGLGRQCPREIGSYNHPDGYTILPG-----LTYMQLQKQDQVDYH 915  
QY 541 LNYNEYIVNPNVNRMYLLKVOFN 565  
DB 916 LLNNEFIVYDQIQLYIVRVKMH 940

RESULT 13  
Q96CG2 ID Q96CG2 PRELIMINARY; PRT; 533 AA.  
AC Q96CG2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 60.1 KDA PROTEIN.  
OS Homo sapiens (Human)  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014260; AAH14260.1; -  
KW Hypothetical protein.  
SQ SEQUENCE -533 AA; 60089 MW; 6296A0B439CC7767 CRC64;

Query Match 24.2%; Score 725; DB 4; Length 533;

Best Local Similarity 35.0%; Pred. No. 1.5e-42;  
Matches 194; Conservative 97; Mismatches 188; Indels 76; Gaps 20;  
QY 46 ESKKMPVAGGKANKDRTEKQDESVALLLKGP-APVDPECTAKVGR-AHYICBGND 100  
DB 15 EKKKGROAG-----REEDPFRSTAEALKAIPAEKRIIRVDPTCLSSNPGTQYI---E 64  
QY 101 VYDMLNQTNLQFNNNKYLIQLLEDQAQRNFVSMRWGRVGMQHSVLVACSNLKA 160  
DB 65 DYNCTLNQINNIENNNKFIYIQLLQ-DSNREFFTCNWRGRVGEVGO-SKINHTRLEDAK 122  
QY 161 EIPQKFLDKTKNWNEDREKFEKVPKGYDMLQMDYATNTODEEETKKEESLKSPLKPESQ 220  
DB 123 KPEKFKREKTKNWAERDHFVSHPGKYTLIEV---QAEDAQEAAYVKVDRGVPVRTVK 178  
QY 221 -----LDLRLVQELIKLCINVOAEMEMMKYNTKKAPLGKLTVAQIKAGYQSLKKIED 274  
DB 179 RVOPCSLDPATQKLTITNIFSKEMFKNTMALMDJVRKMPGLKLSKQOIARGFFALEALEE 238  
QY 275 CIRA-GQHGRLMEACNEFYTRIPHDGFLRTPLPRTQKELSEKIQLEALGDIEI--AI 331  
DB 239 ALKGPDTGGGSLSELSHFYVIPHNFGHSQPPPIINSPELLQAKKMDLLVLADIQAAL 298  
QY 332 KLVTTELQSP-----HPLDQHYRNLCALRPLDHESEYFKVISQYLOSTHAPTHSDYTM 388  
DB 299 QAYSEQEKTVVEYPHPLDRDYQLLKCQLQDLSGAPEKYVIQYLEQGTGS---NHRCP 355  
QY 389 LDLFVEVEKDGKEAFAFR-EDLHNRMLLWHGSRMNWVGILSHGLRTAPPEAITGYMGFK 446  
DB 356 QHTWKVNGQEGEDRFQAHSKLGNRKLWHGTNMAVNAIITSLGRIMPH-----SGR 411  
QY 447 GIYFADMSKKSANYCFASR--LKNTGLLLSEVALGQCNEELLEANPKAEGLLQGHSTKG 504  
DB 412 GIYFASSENSKAGYVIGMKCGAHVGMFLGEALGREGHINTDNPSL----- 459  
QY 505 LGMARSSAHFVTLNGSTVPLGPASDT-----GILNPDGY-----TLNNEYI 547  
DB 460 --KSPPPGFDVITARGHTEP-DPTQDTELELDGQVVVPOGPVPCPFSSSTFSQSEYL 516  
QY 548 VYNPNVNRMYLLKLV 562  
DB 517 IYVESQCRRLYLEV 531

RESULT 14  
Q91YR6 ID Q91YR6 PRELIMINARY; PRT; 528 AA.  
AC Q91YR6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 59.4 KDA PROTEIN.  
OS Mus musculus (Mouse)  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014870; AAH14870.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 528 AA; 59413 MW; 82EPB0C498EB5F74 CRC64;

Query Match 23.2%; Score 696.5; DB 11; Length 528;  
Best Local Similarity 34.5%; Pred. No. 1.4e-40;  
Matches 184; Conservative 92; Mismatches 213; Indels 45; Gaps 17;  
QY 54 GSKANKDRTEKQDESVALLLKGP-----VDPECTAKVKAHYICBGNDVYDML 106  
DB 13 GSKKQKQGTEDSFRSTAEALRA-APADNRVIRVDPCFSPSRNPG---IQVHEDYDCTL 68

```

QY 107 NOTNLOFNKKYIIQLLEDDAQRNFVVMWRGVRGKMGQHSI--VACSGNLNKAKEIFQ 164
    ||||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 69 NOTNIGNNNKKYIIQLLEGSR--FFCWRNRGRVGEVQSGKNHFTC---LEDAKDKFK 123
    ||||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 165 KFLDKTKNNWDEKFEKVPKQYDMLQMDYATNTQDEETKEESLSKPLKPESOLDLR 224
    ||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 124 KFWFKTKNKNWEERDFVAQPNKYITLIEVQGEAESQEAIVKVDSPVRTVWVKPCS-LDPA 182
    ||||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 225 VQELIKLIGNVOAMEEMMEMYNTKKAPLGKLTVAQIKAGYOSLKKIEDCIR--AGQHG 282
    ||||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 183 TQNLITNIFSKEMFRNAMTLMVDMVKMPLGKLTQOIARGFEALEEAMKNPTGD-G 241
    ||||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 283 RALMEACNEFYTRIPHDGFLRTPPLRTQKELSEKTLQLEALGDIEIAIKL-----VKT 336
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 242 QSLSELSFCFYVIPHNFGRSRPPPIINSPDVQAKDKMLLVADIELVOTLQAAPEEREE 301
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 337 ELQSPHPDLQHYRNHLCALRPLDHESYEFKVISQYLOSTHAPTHSDYTMTLDLFEVBK 396
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 302 KVEEVPHPDLRDYQLLRCOLLDSESEYKAIQTYLKQT--GNSYRCPNLRHVKVNR 358
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 397 DGEKEAFR--EDLHNRMLLWHGSRMWNWYILSHGLRIAPPEAPITGYMFGKIYFADMS 454
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 359 EGGEDRFQASHKLGNNRLLWHGTNVAVAAIITSLGRIMPH----SGGRVGKIYFASEN 414
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 455 SKSANYCFASRL--KNTGILLSEVALGOCNELLEANPKAEGILQGHSTKGLGKMAPSS 512
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 415 SKSAGVTTTHCGGHQGVYMFGEVALGKEHHTIDDPKSPPGFDSVIARGQTEPDP 474
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 513 AHFV--TLNGS--TVPLGPASTGILNPDGYTLNNEYIVYNPNQVRYMLLKV 562
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 475 AODIELELDQPVVVPQPPVQPCPSFKSSF--SQSEYLIYKESQCRRLYLEI 526
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:

RESULT 15
QXUA5 PRELIMINARY; PRT; 727 AA.
AC Q9XUA5;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE AC8.1 PROTEIN.
GN AC8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
[1]
RN SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; 283097; CAB05448.1; -.
DR HSSP; P26446; IAZ6.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; Znf-PARP; 1.
DR ProDom; PD004675; Znf-PARP; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFEEF CRC64;

```

Query Match 22.9%; Score 685.5; DB 5; Length 727;  
 Best Local Similarity 31.1%; Pred. No. 1.3e-39;  
 Matches 178; Conservative 110; Mismatches 220; Indels 65; Gaps 16;

```

QY 19 ESKRVNNGNTAPEDSSPAKTRRCORQESKMPVAGGKANKDRTEDKQDESVALLLKG 78
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 189 BALAAKGGSTFATPASASPT-----PPEAETPVLSAEGSPE--SSNKRPSAEIIDGE 242
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 79 APVDPPECTAK-----VGKAHVYCEGNDVYDVMNLNOTNLOFNKKYIIQLLED 126
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 243 GNPDENDFAKKRMKKEARLMEVQKRMKKQSDLLWEYRQIPERMPYTON----ISILRE 298
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 127 DAQRNFVVMWRGVRGKMGQHSI--VACSGNLNKAKEIFQKFLDKTKNNWDEKFEKVP 186
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 299 NEQD-----IPEGHRTAQDYFH-----EKTNDWIYRKHFRKMPG 334
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 187 KYDMLQMDYATNTQDEETKKEESLSKPLKPESO--LDLRVQELIKLIGNVOAMEEMME 244
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 335 MFSYVETDISFEVGTNNGHKKITPGSKTLLPKSVKEVYVMSITFDVENMKSAKLS 394
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 245 MKYNTKAPLGKLTVAQIKAGYOSLKKIED-CIRAGQHGGRALMEACNEFYTRIPHDGFLR 303
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 395 FEIDVNMPLGRLSHNQINLAPEVLNDISDLLVKLPIDASKILDFSNKFVTIIPHNFGMR 454
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 304 TPPLRTQKELSEKTLQLEALGDIEIAIKV-----KTELQSPHPDLQHYRNHLCAL 356
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 455 VPEPIDSFHKIKKNNMLNALLDIFKPAYDQISGDDVPASTLSI--DPVDINRYKLKCI 512
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 357 RPLDHESYEFKVISQYLOSTHAPTHSDYTMTLDLFEVBKDGKEAFREDLHNRMLLWHG 416
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 513 EPLQCCDDNMHIOYKLNTHGATH--DLKVELIDILKVNHDNESSFKRKHIGNRLLWHG 571
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 417 SRMSNMGVILSHGLRIAPPEAPITGYMFGKIYFADMSKSKSANYCFASRLKNTGLLLSE 476
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 572 SGKMNFAgilGQGLRIAPPEAPVSGYMGKGVYFADMFMSKSPFYCRAN-AKEEAYLLLCD 630
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 477 VALGQCNELLEA-NPKAEGILQGHSTKGLGKMAPSS-AHFVTLNGSTVPLGPA--SDTG 532
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 631 VALGNVQQLMASKNVSRQTLPAQFQSVQGVGRCQPREIGSYKNPDGYTVPLGLTYMLOG 690
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
QY 533 ILNPDGYTLNNEYIVYNPNQVRYMLLKVQFN 565
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:
Db 691 QNVVD-YHLLYNEFIVYDQIQLKYLVRVKMH 722
    :||| :|||:|||||: :| :|||:|||||: :| :|||:|||||: :| :|||:|||||:

```

Search completed: August 29, 2002, 08:01:28  
 Job time: 360 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:57:12 ; Search time 124.84 Seconds  
(without alignments)  
474.226 Million cell updates/sec

Title: US-09-701-586B-4  
Perfect score: 2823  
Sequence: 1 MAPKPKPWQTEGPEKKGR.....EVLIIQESQCRRLRLLEVLH 533

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2823	100.0	533	21	AAV511175 Human brain PARP3
2	2823	100.0	540	21	AAV511176 Human uterus type
3	2811	99.6	533	22	AAU29021 Human PARP-3 prote
4	2253.5	79.8	528	21	AAV511178 Murine PARP1 (shor
5	2241	79.4	533	21	AAV511177 Murine PARP1 (long
6	738	26.1	522	22	AAU29022 Mouse PARP-2 prote
7	729.5	25.8	583	22	AAB47029 hPARP2. Homo sapi
8	726.5	25.7	534	21	AAB42909 Human ORFX ORF2673
9	726.5	25.7	534	22	AAU29023 Human PARP-2 prote
10	725.5	25.7	534	22	AAU29020 Human PARP-2 prote
11	725	25.7	570	21	AAV511174 Human brain PARP2

12	725	25.7	570	22	AAB11480 Human brain poly-A
13	722	25.6	521	22	AAB60693 Human poly(ADP-rib
14	688.5	24.4	637	21	AAB68835 The poly(ADP-ribos
15	682.5	24.2	531	22	AAV93513 Human protein sequ
16	670.5	23.8	653	21	AAV68834 A poly(ADP-ribose)
17	668.5	23.7	1014	21	AAV58043 Human poly (ADP-ri
18	668.5	23.7	1014	22	AAU29019 Human PARP-1 prote
19	668.5	23.7	1014	22	AAV66296 Human tankyrase2 r
20	667.5	23.6	1014	21	AAV49939 Human nuclear NAD+
21	666.5	23.6	1013	17	Poly(ADP-ribose) p
22	660.5	23.4	1014	20	Human poly(ADP-rib
23	638	22.6	557	22	Drosophila melanog
24	637	22.6	1063	22	AAB47032 Fusion protein PAR
25	578	20.5	969	21	A poly(ADP-ribose)
26	578	20.5	980	21	A poly(ADP-ribose)
27	475	16.8	982	20	Maize poly ADP-rib
28	472.5	16.7	379	22	Novel human neopla
29	447	15.8	360	22	N-terminal fragmen
30	410.5	14.5	1010	21	Fusion protein of
31	394	14.0	294	22	Human DNA repair a
32	394	14.0	294	22	Novel human neopla
33	390	13.8	287	22	C-terminal fragmen
34	284.5	10.1	1099	22	Human tankyrase2 e
35	267.5	9.5	227	22	Human DNA repair a
36	267.5	9.5	227	22	Novel human neopla
37	258	9.1	1730	22	Novel human secret
38	256	9.1	1724	21	CDNA sequence enco
39	256	9.1	1724	22	Human minor vault
40	216.5	7.7	190	22	Novel human neopla
41	138.5	4.9	1327	21	Human tankyrase I
42	138.5	4.9	1327	21	Human tankyrase
43	138.5	4.9	1327	22	Human tankyrase1 S
44	137	4.9	756	22	Human tankyrase2 C
45	137	4.9	784	22	Human tankyrase2 C

ALIGNMENTS

RESULT 1

AAV511175  
ID AAV511175 standard; Protein: 533 AA.

AC AAV511175;

DT 31-MAR-2000 (first entry)

XX Human brain PARP3 protein.

DE PARP: poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP3.

OS Homo sapiens.

XX WO9964572-A2.

XX 16-DEC-1999.

PF 04-JUN-1999; 99WO-EP03889.

PR 05-JUN-1998; 98DE-1025213.

PR 01-MAR-1999; 99DE-1008837.

(BADI ) BASF AG.

XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;

DR WPI; 2000-087218/07.

DR N-ESDB; AAZ44288.

PT Novel genes and proteins, antibodies and binding partners useful in  
diagnosis and therapy of energy deficiency associated disease

PT conditions -  
PS Claim 4; Page 57-59; 96pp; German.  
XX  
XX This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>C (I). The nucleic acid  
CC sequences, PARP homologues and antibodies are useful for analytic  
CC detection of PARP homologues and for identifying PARP effectors or  
CC binding partners, as well as for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
CC human PARP3 protein used in the method of the invention.  
XX  
XX Sequence 533 AA;  
XX  
Query Match 100.0%; Score 2823; DB 21; Length 533;  
Best Local Similarity 100.0%; Pred. No. 1.1e-249;  
Matches 533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPKPKWVQTEGPEKKKGQAGREEDPFRSTAEALKAIPAERIRVDPCTPLSSNPGT 60  
DB 1 mapkpkpwvqtgpekkkgqagreedpfrstaealkaipaerirrvdptcplssnpgt 60  
QY 61 QVYEDYNTLQNIENNNKFIYIQLQDSNRRFFTCWNRWGRVGEVQSGKINHFRLED 120  
DB 61 qvyedynctlnqniennnkfyiqlldsnrfftcwnrwwgrvgevqsgskinhfrled 120  
QY 121 AKDFFKFKREKTKNNWAERDHFVSHPGKYTLIEVQAEQAQVAVKVDGPGVTVTKRV 180  
DB 121 akdfeffkfrcktannwaerdhfvshpgktyllievdaeqaavvkdrgpvtvtkrv 180  
QY 181 QPCSLDPATOKLTNTIFSKEMFNKTMALMDLVKMPGLGKLSKQQTARGFEALEALEAL 240  
DB 181 qpcslpdatokltnifskemfntmalmdlvkmpglgklskqqiargfealealeaal 240  
QY 241 KGPTDGGQSLSELSHFYTVIPHNFGSQPPPINSPPELLQAKKMDLLVLADIELAQALQA 300  
DB 241 kgptdggqslseelsshfytviphnfgsqpppinpspellqakkmdllvladielaqalqa 300  
QY 301 VSEQEKTVVEVPPLDRDYOLLKCOLQLDLSGAPEYKVIQTYLEQTSNHRCPDLOHIWK 360  
DB 301 vseqektvveevphpldrdyollkcolqlldsgapeykvityleqtsgnhrcptlqhniwk 360  
QY 361 VNOEGEEDRFQAHKSLGNRRKLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGIVFASENS 420  
DB 361 vnogeedrfqahkslgnrrkllwhgtnmavvaailtsglrimphsgrgvgkgivfasens 420  
QY 421 KSGAGYVIGMKCGAHVGYMFLGEVALGREHHINTDNPDLKSPPPGDSVITARGHTPEDPPT 480  
DB 421 ksagayvigmkcgahvgymlgevalgrehhintdnpdlkspppgdsvitarghtepdpt 480  
QY 481 QDTELELDGQGVVPOGVPVPCPEFSSTFESQSEYLYIQESQCLRLRYLLEVHL 533  
DB 481 qdteleldgqgvvpggvpcpcefsstfsesqseyllyiqesqclryllevgl 533  
RESULT 2  
ID AAY51176  
XX AAY51176 standard; Protein; 540 AA.  
AC AAY51176;  
XX  
XX 31-MAR-2000 (first entry)  
DT  
XX Human uterus type 2 PARP3 protein.  
DE  
XX

KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP3.  
XX Homo sapiens.  
XX OS  
XX W09964572-A2.  
XX PN  
XX 16-DEC-1999.  
XX PD  
XX 04-JUN-1999; 99WO-EP03889.  
XX PF  
XX 05-JUN-1998; 98DE-1025213.  
XX PR  
XX 01-MAR-1999; 99DE-1008837.  
XX PP  
XX (BADI ) BASF AG.  
XX PI  
XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX WPI; 2000-087218/07.  
XX DR  
XX N-PSDB; AA244289.  
XX  
XX Novel genes and proteins, antibodies and binding partners useful in  
XX diagnosis and therapy of energy deficiency associated disease  
XX conditions -  
XX  
XX Claim 4; Page 62-64; 96pp; German.  
XX  
XX This invention describes novel human and murine poly(ADP-ribose)  
XX polymerase (PARP) homologues, which are characterised by an amino acid  
XX sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
XX sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>C (I). The nucleic acid  
XX sequences, PARP homologues and antibodies are useful for analytic  
XX detection of PARP homologues and for identifying PARP effectors or  
XX binding partners, as well as for determining their effectiveness.  
XX PARP-binding partners are useful for the diagnosis or therapy of a  
XX disease condition, which is the result of a PARP protein, especially an  
XX energy deficiency, which may comprise tissue damage from cell death  
XX following necrosis or apoptosis. The disease condition may be chosen  
XX from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
XX in particular neurotoxic disturbances, etc. This sequence represents the  
XX human PARP3 protein used in the method of the invention.  
XX  
XX Sequence 540 AA;  
XX  
Query Match 100.0%; Score 2823; DB 21; Length 540;  
Best Local Similarity 100.0%; Pred. No. 1.1e-249;  
Matches 533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MAPKPKWVQTEGPEKKKGQAGREEDPFRSTAEALKAIPAERIRVDPCTPLSSNPGT 60  
DB 8 mapkpkpwvqtgpekkkgqagreedpfrstaealkaipaerirrvdptcplssnpgt 67  
QY 61 QVYEDYNTLQNIENNNKFIYIQLQDSNRRFFTCWNRWGRVGEVQSGKINHFRLED 120  
DB 68 qvyedynctlnqniennnkfyiqlldsnrfftcwnrwwgrvgevqsgskinhfrled 127  
QY 121 AKDFFKFKREKTKNNWAERDHFVSHPGKYTLIEVQAEQAQVAVKVDGPGVTVTKRV 180  
DB 128 akdfeffkfrcktannwaerdhfvshpgktyllievdaeqaavvkdrgpvtvtkrv 187  
QY 181 QPCSLDPATOKLTNTIFSKEMFNKTMALMDLVKMPGLGKLSKQQTARGFEALEALEAL 240  
DB 188 qpcslpdatokltnifskemfntmalmdlvkmpglgklskqqiargfealealeaal 247  
QY 241 KGPTDGGQSLSELSHFYTVIPHNFGSQPPPINSPPELLQAKKMDLLVLADIELAQALQA 300  
DB 248 kgptdggqslseelsshfytviphnfgsqpppinpspellqakkmdllvladielaqalqa 307  
QY 301 VSEQEKTVVEVPPLDRDYOLLKCOLQLDLSGAPEYKVIQTYLEQTSNHRCPDLOHIWK 360  
DB 308 vseqektvveevphpldrdyollkcolqlldsgapeykvityleqtsgnhrcptlqhniwk 367



QY 361 VNOQGEEDRFOAHSKLGKRNKLLWHTGNMVAAILTSGLRIMPHSGRGVKGKIYFASENS 420  
|||||  
Db 368 vnqgeedrfqahsklgnrkllwhtgnmavvaailtsglrinphsggrvgkiyfasens 427  
QY 421 KSAGYVIGMKCGAHVGYMFLGEVALGRHHINTDNPSLKSPPGPDSDVIARGHTEPDPT 480  
Db 428 ksagyyigmkcgahvgymflgevalgrehhintdnpslkspppgfsviarghtepdpt 487  
QY 481 QTELELDGQQVVPQGPVPCPEFSSSTFSQSEVLIYQESQCRRLYLEVHL 533  
|||||  
Db 488 qdteleldgqqvvpqgvpvpcpefssstfsqseyllyqesqcrlyllevhl 540

RESULT 3  
AAU29021  
ID AAU29021 standard; Protein; 533 AA.  
XX  
AC AAU29021;  
DT 18-DEC-2001 (first entry)  
DE Human PARP-3 protein.  
XX  
KW Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
KW cytosolic; nootropic; neuroprotective; antiinflammatory; antidiabetic;  
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
XX  
OS Homo sapiens.  
XX  
PN WO200164955-A1.  
XX  
PD 07-SEP-2001.  
XX  
PF 01-MAR-2001; 2001WO-US06572.  
XX  
PR 02-MAR-2000; 2000US-0517467.  
XX  
PA (ISIS-) ISIS PHARM INC.  
PI Popoff I, Cowser LM;  
XX  
XX WPI; 2001-602570/68.  
DR N-PSDB; AAS45590.  
XX  
XX Antisense compound useful for treating hyperproliferative,  
PT neurological, inflammatory and autoimmune disorders and diabetes  
PT inhibits human PARP -  
XX  
PS Example 13; Page 105-107; 168pp; English.  
XX

The invention relates to antisense oligonucleotides targeted to human  
CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, DNA repair, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARP in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARP especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
CC arthritis) and diabetes. The present sequence is a PARP protein,  
CC the cDNA encoding which was used to design the antisense  
CC oligonucleotides.  
XX  
SQ Sequence .533 AA;

Best Local Similarity 99.6%; Pred. No. 1.3e-248;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAPKPKPWQVTGPEKKKGKQAGREDPPRSTAEALKAIPAERIRIVDPPTCLSSNPOT 60  
|||||  
Db 1 mapkpkpwvqtgpekkkgkrgreepfrstaealkaipaekriirvdpptclssnpot 60  
QY 61 QVYEDYNCTLNQNTENNKNFYIIQLQDSNRFFTCNNRWGRVGVGSGKINHFTRLSD 120  
|||||  
Db 61 qvyedynctlnqntennknkfyiqlqdsnrfftcnnrwgrvgevsgskinhfrl 120  
QY 121 AKKDEKFKPREKTKNWAERDHFVSHPGKTYLIEVOADEAQEAVVYVDRGVRVTVKRV 180  
|||||  
Db 121 akkdfekfkprektknwaerdhfvsphgktylievoadeaqeavvkvdrapvrtvkrv 180  
QY 181 QPCSILDPATOKLITNIFSKEMFKNTMALMDLVKKMPLGKLSKQQTARGFEALEEAL 240  
|||||  
Db 181 qpcsilpatoklitnifskemfkntmalmdldvkkmpigklskqqiargfealeeal 240  
QY 241 KGPTDGGQSLEELSSHFYTVIPHNFGHSOPPPINSPPELLQAKKMLVLADIQAALQA 300  
|||||  
Db 241 kgptdggqsleelsshfytviphnfghspppinpellqakkmllvdielaqalqa 300  
QY 301 VSEQKTVVEVPHPLDRDYQLLKQQLDLSGAPEYKVIQTYLEQTSNHRCPPTLQHIWK 360  
|||||  
Db 301 vseqktveevphpldrdyqllikcqlldlsgapekviqtyleqtsnhrctqhiwk 360  
QY 361 VNOQGEEDRFOAHSKLGKRNKLLWHTGNMVAAILTSGLRIMPHSGRGVKGKIYFASENS 420  
|||||  
Db 361 vnqgeedrfqahsklgnrkllwhtgnmavvaailtsglrinphsggrvgkiyfasens 420  
QY 421 KSAGYVIGMKCGAHVGYMFLGEVALGRHHINTDNPSLKSPPGPDSDVIARGHTEPDPT 480  
|||||  
Db 421 ksagyyigmkcgahvgymflgevalgrehhintdnpslksppgfdsviarghtepdpt 480  
QY 481 QTELELDGQQVVPQGPVPCPEFSSSTFSQSEVLIYQESQCRRLYLEVHL 533  
|||||  
Db 481 qdteleldgqqvvpqgvpvpcpefssstfsqseyllyqesqcrlyllevhl 533

RESULT 4  
AA51178  
ID AA51178 standard; Protein; 528 AA.  
XX  
AC AA51178;  
XX  
DT 31-MAR-2000 (first entry)  
XX  
DE Murine PARP1 (short) homologue protein.  
XX  
KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP1.  
XX  
OS Mus sp.  
XX  
PN WO9964572-A2.  
XX  
PD 16-DEC-1999.  
XX  
PF 04-JUN-1999; 99WO-EP03889.  
XX  
PR 05-JUN-1998; 98DE-1025213.  
PR 01-MAR-1999; 99DE-1008837.  
XX  
PA (BADI ) BASF AG.  
XX  
PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX WPI; 2000-087218/07.  
DR N-PSDB; AAZ44291.  
XX  
PT Novel genes and proteins, antibodies and binding partners useful in

PT diagnosis and therapy of energy deficiency associated disease  
 PT conditions -

PS Claim 4; Page 71-73; 96pp; German.

XX This invention describes novel human and murine poly(ADP-ribose)  
 CC polymerase (PARP) homologues, which are characterised by an amino acid  
 CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
 CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>2C (I). The nucleic acid  
 CC sequences, PARP homologues and antibodies are useful for analytic  
 CC detection of PARP homologues and for identifying PARP effectors or  
 CC binding partners, as well as for determining their effectiveness.  
 CC PARP-binding partners are useful for the diagnosis or therapy of a  
 CC disease condition, which is the result of a PARP protein, especially an  
 CC energy deficiency, which may comprise tissue damage from cell death  
 CC following necrosis or apoptosis. The disease condition may be chosen  
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
 CC in particular neurotoxic disturbances, etc. This sequence represents the  
 CC murine PARP1 protein used in the method of the invention.

XX Sequence 528 AA;

Query Match 79.8%; Score 2253.5; DB 21; Length 528;  
 Best Local Similarity 80.9%; Pred. No. 1.7e-197;  
 Matches 432; Conservative 37; Mismatches 58; Indels 7; Gaps 4;

QY 1 MAPKPKVQTEGPEKKKGROAGREDFRSTAEALKAIPAEEKRIIRVDTPCLSSNPGT 60  
 DB 1 mapkrkasvqtg--skkqrqgteedsfrstaealraapadnrvirvdpfcfsrnpqi 58  
 QY 61 QVYEDYNTLQNTNIENNKKFYIIQLQDSNRFCTCNRWGRVGEVQSKINHTFRLD 120  
 DB 59 qvhedyctlnqtnignnnkfyiiqlleesrff-cwnrgrvgrvgeqskmhfctled 117  
 QY 121 AKDPEKFKREKTKNNNAERDHFVSHPGKYTLIEVQAEDEAQAQAVKVDGPGVTRTV 180  
 DB 118 akdkfkkfwektnkweerdfrvaqpnkytlievqgeaesqaavkvdsqgvrtv--v 174  
 QY 181 QPCSLDPATQKLTNIFSKEMFKNTMALDVKKMPGLKSLKQOIARGFEALEAL 240  
 DB 175 kpcslpatqnlitnifskemfknamtlnmldvkkmpglkltqgiargfealealeam 234  
 QY 241 KGPTDGGQSLEELSHFYTVIPHNFGHSQPPPIINSPPELLQAQKMDLLVLADI 300  
 DB 235 knptgdgqslleelsscfytlphnfgsrpppnsdpvlqakkmllvladielaqltqa 294  
 QY 301 V-SEQEKTVEVPPLDRDYQLLCKQLDLSGAPEYKVIQTYLEQTSNHRCP 359  
 DB 295 apgeeeekveevphldrdyqlrcqlldsgeseykaiqlylkgtgnsyrcpnlrhv 354  
 QY 360 KVNQGEDEDFQAHSLGNRKLWHGTNMAVAAILTSLRIMPESHGGRVKGIFYFAS 419  
 DB 355 kvnregeedrfqahsklgnrrllwhgtnavvaailltsglrimpshggrvkgifyfas 414  
 QY 420 SKSAGYVIGMKCAHGYMFLGVALGREHHINTDNPSLSPPPGSDSVIARGHTPE 479  
 DB 415 sksagvvtmhcgghqymflgvalgkhehhtidpdlkspppgfsvdiargtqepd 474  
 QY 480 TDTELELDGQVVVPGQVPCPEFSSFSQSEYLIYQESQRLRYLLEVHL 533  
 DB 475 aqdieleldgqvvpvpgpvpcpsfssfsqseyllykesqrlryllehl 528

RESULT 5

AA51177 standard; Protein; 533 AA.

XX  
 AC AA51177;

DT 31-MAR-2000 (first entry)

XX Murine PARP1 (long) homologue protein.

XX PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
 KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
 KW ischemic tissue damage; PARP1.

XX Mus sp.

XX WO9964572-A2.

XX 16-DEC-1999.

XX 04-JUN-1999; 99WO-EP03889.

XX 05-JUN-1998; 98DE-1025213.

XX 01-MAR-1999; 99DE-1008837.

XX (BADI ) BASF AG.

XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;

XX WPI: 2000-087218/07.

XX N-PSDB; AA244290.

XX Novel genes and proteins, antibodies and binding partners useful in  
 PT diagnosis and therapy of energy deficiency associated disease  
 PT conditions -

XX Claim 4; Page 67-69; 96pp; German.

XX This invention describes novel human and murine poly(ADP-ribose)  
 CC polymerase (PARP) homologues, which are characterised by an amino acid  
 CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
 CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>2C (I). The nucleic acid  
 CC sequences, PARP homologues and antibodies are useful for analytic  
 CC detection of PARP homologues and for identifying PARP effectors or  
 CC binding partners, as well as for determining their effectiveness.  
 CC PARP-binding partners are useful for the diagnosis or therapy of a  
 CC disease condition, which is the result of a PARP protein, especially an  
 CC energy deficiency, which may comprise tissue damage from cell death  
 CC following necrosis or apoptosis. The disease condition may be chosen  
 CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
 CC in particular neurotoxic disturbances, etc. This sequence represents the  
 CC murine PARP1 protein used in the method of the invention.

XX Sequence 533 AA;

Query Match 79.4%; Score 2241; DB 21; Length 533;  
 Best Local Similarity 80.1%; Pred. No. 2.4e-196;  
 Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

QY 1 MAPKPKVQTEGPEKKKGROAGREDFRSTAEALKAIPAEEKRIIRVDTPCLSSNPGT 60  
 DB 1 mapkrkasvqtg--skkqrqgteedsfrstaealraapadnrvirvdpfcfsrnpqi 58  
 QY 61 QVYEDYNTLQNTNIENNKKFYIIQLQDSNRFCTCNRWGRVGEVQSKINHTFRLD 120  
 DB 59 qvhedyctlnqtnignnnkfyiiqlleesrff-cwnrgrvgrvgeqskmhfctled 117  
 QY 121 AKDPEKFKREKTKNNNAERDHFVSHPGKYTLIEVQAEDEAQAQAVK----VDRGPV 175  
 DB 118 akdkfkkfwektnkweerdfrvaqpnkytlievqgeaesqaavkvdsqgvrt 177  
 QY 176 VTRVQPCSLDPATQKLTNIFSKEMFKNTMALDVKKMPGLKSLKQOIARGFEALEA 235  
 DB 178 v---vkpcslpatqnlitnifskemfknamtlnmldvkkmpglkltqgiargfealea 234  
 QY 236 LEEALKGPTDGGQSLEELSHFYTVIPHNFGHSQPPPIINSPPELLQAQKMDLLVLADI 295  
 DB 235 leeamkntgdgqslleelsscfytlphnfgsrpppnsdpvlqakkmllvladiela 294  
 QY 296 QALQAV-SEQEKTVEVPPLDRDYQLLCKQLDLSGAPEYKVIQTYLEQTSNHRCP 354  
 DB 295 aqdieleldgqvvpvpgpvpcpsfssfsqseyllykesqrlryllehl 528

Db 295 qtlqaapgeeeekveephpldrdyqlrcqlqldsgeseykaiqltylktgnsyrcpn 354  
QY 355 LOHIWVNOGEBEDRFQAHSKLGNRRLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGYY 414  
Db 355 lrhwkvnrgegedrfqahsklgnrrllwhgtnavvaailtesglrhmshggrvgkgiy 414  
QY 415 FASENSKSGYVIGMKCGAHVGYMFLGEVALGREHHINTDNP SLKSPPPGFDVSIARGH 474  
Db 415 fasensksagvytmhcgghqvgymflgevalgkehhtiddpslkspppgfdsviargg 474  
QY 475 TEPDPTQDTELELDGQOVVVGOPVPCPEFSSSTFQSEYLLIYQESQCLRLYLLVHL 533  
Db 475 tepdpqdieleldgqovvvgppvqcpbsfkssfsqseyllykesqrllylleihl 533  
RESULT 6  
ID AAU29022 standard; Protein; 522 AA.  
XX AAU29022;  
AC AAU29022;  
DT 18-DEC-2001 (first entry)  
DE Mouse PARP-2 protein.  
XX Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
KW cytostatic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;  
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
OS Mus musculus.  
XX WO200164955-A1.  
PN 07-SEP-2001.  
XX 01-MAR-2001; 2001WO-US06572.  
PF 02-MAR-2000; 2000US-0517467.  
PR (ISIS-) ISIS PHARM INC.  
XX Popoff I, Cowser LM;  
XX WPI; 2001-602570/68.  
DR N-PSDB; AAS45397.  
XX Antisense compound useful for treating hyperproliferative,  
PT neurological, inflammatory and autoimmune disorders and diabetes  
PT inhibits human PARP -  
XX Example 13; Page 109-111; 168pp; English.  
XX The invention relates to antisense oligonucleotides targeted to human  
CC PARY nucleic acid and inhibiting expression of human PARY. PARY  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, DNA repair, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARY in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARY especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
CC arthritis) and diabetes. The present sequence is a PARY protein,  
CC the cDNA encoding which was used to design the antisense  
CC oligonucleotides.  
XX Sequence 522 AA;  
SQ

Query Match 26.1%; Score 738; DB 22; Length 522;  
Best Local Similarity 34.3%; Pred. No. 1.5e-58;  
Matches 187; Conservative 103; Mismatches 191; Indels 64; Gaps 19;  
QY 16 KKGROA-CREDDPPFRSTAEALKAIPAEKRIIRVDTCP LSSNPGTQVY-----EDYNCIL 70  
Db 5 qrkqpmagdkdadrtdkndrdsvktlllkdk-apvdecaaklkg-ahvycegdvdydvm 62  
QY 71 NOTINNNNNKFVLIQLO-DSNRFFTCNWRGRCVGEVQ-SKINHFTLEDAKKDFEKK 128  
Db 63 nqtnlqfnunkyyllqldeddaqrnfsvmrgvrgvgtqgshlvtscgdlnkakeifqk 122  
QY 129 FREKTNNNAERDHFVSHPGKYTLIEV-----QAEDPA---QEA VVKVDRGPVRTVTKRVO 181  
Db 123 fldktnnwedrenfekvpgkydmlqmdyaastqdesktkeetlkpe----- 170  
QY 182 PCSLDPATOKLITNTPSKEMFNTMALMDLVKMKPLGLSKLQKQIARGFEALEEALK 241  
Db 171 -sqldlrqvdekllicnvqvtmeemmiemkydtkraplgtkltvaqikagvgsllkiedcir 229  
QY 242 GPTDGGOSLEELSSHFYTVIPHNGHSQPPINSPPELLQAKKMDMLVLADIELAQALQAV 301  
Db 230 a-qghgralveacnefytriphdfglisppvirtekelsdkvkllealgdiei--alkiv 286  
QY 302 SEQKTEVEVPHPDLRDYQLKCOLQLDLSGAPEYKVIQTYLEQT-GSNHR--CPTLQHI 358  
Db 287 kserqgle--hpldqhyrnlhcalrpldhesnefkvisqyiqsthapthkdytmlldv 343  
QY 359 WKVNBQGEEDRQANSKLNKRLWHGTNMAVVAAILTSGLRIMPH-----SGRGVKGYY 414  
Db 344 fevekegekeaf--edlpnrmlwhgslrslwvngilshglrvappeapitgymfgkgy 401  
QY 415 FASENSKSGYVIGMKCGAHVGYMFLGEVALGREHHINTDNP SLKSPPPGFDVSIARGH 474  
Db 402 fadmsksanycfasr--lkntglillilsevalgcnelleanpkadgllrghstkgm 459  
QY 475 TEPDPTQDTELELDGQOVVVGOPVPCPEFSSS-----TFQSEYLLIYQESQCLRL 526  
Db 460 mapsa-----hfitngstvpplpasdtglnpgeytlvnynefivspnqvmr 509  
QY 527 YLLEV 531  
Db 510 yllki 514  
RESULT 7  
AAB47029  
ID AAB47029 standard; Protein; 583 AA.  
XX AAB47029;  
AC AAB47029;  
DT 29-MAR-2001 (first entry)  
XX 29-MAR-2001 (first entry)  
DE hPARP2.  
XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;  
KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;  
KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;  
KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;  
KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;  
KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;  
KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;  
KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;  
KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;  
KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;  
KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;  
KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;  
KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;  
KW graft versus host disease; allograft rejection; cystic fibrosis;  
KW chronic glomerulonephritis; inflammatory bowel disease;  
KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;  
KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;  
KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;







```

RESULT 11
AA51174
ID AAV51174 standard; Protein; 570 AA.
XX
AC AAV51174;
XX
DT 31-MAR-2000 (first entry)
XX
DE Human brain PARP2 protein.
XX
KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW ischemic tissue damage.
XX
OS Homo sapiens.
XX
PN WO9964572-A2.
XX
PD 16-DEC-1999.
XX
PF 04-JUN-1999; 99WO-EP03889.
XX
PR 05-JUN-1998; 98DE-1025213.
PR 01-MAR-1999; 99DE-1008837.
XX
PA (BADI ) BASF AG.
XX
PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR WPI; 2000-087218/07.
DR N-PSDB; AAZ44287.
XX
PT Novel genes and proteins, antibodies and binding partners useful in
PT diagnosis and therapy of energy deficiency associated disease
PT conditions -
XX
PS Claim 4; Page 52-54; 96pp; German.
XX
CC This invention describes novel human and murine poly(ADP-ribose)
CC polymerase (PARP) homologues, which are characterised by an amino acid
CC sequence with a functional NAD+-binding site and no zinc finger
CC sequence motif, of general formula CX2CX2MX2 (1). The nucleic acid
CC sequences, PARP homologues and antibodies are useful for analytic
CC detection of PARP homologues and for identifying PARP effectors or
CC binding partners, as well as for determining their effectiveness.
CC PARP-binding partners are useful for the diagnosis or therapy of a
CC disease condition, which is the result of a PARP protein, especially an
CC energy deficiency, which may comprise tissue damage from cell death
CC following necrosis or apoptosis. The disease condition may be chosen
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,
CC in particular neurotoxic disturbances, etc. This sequence represents the
CC human PARP2 protein used in the method of the invention.
XX
SQ Sequence 570 AA;

Query Match 25.7%; Score 725; DB 21; Length 570;
Best local Similarity 35.4%; Pred. No. 2.6e-57;
Matches 197; Conservative 95; Mismatches 185; Indels 80; Gaps 21;

Qy 15 EKKKGKQAG-----REEDPFSTAEALKAPEAKRIIRVDPTCPLSSNPG-TQVY---- 63
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
46 eskmpvaggkankdrted---kqdesvkalllkgk-apvdec--takvgkahvcegn 99

Qy 64 EDYNTLNTQNIENNKNFYIQLIQ-DSNRFTTCWNKRWGEVQ-SKINHFTPLRLEDA 121
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
100 dyvdvmlnqtnlqfnnkyyliqlledaqnfnswmrgvrgknqgshlvacsnglnlka 159

Qy 122 KDFEKKFREKTKNNWAERDFVSHPGKYTLIEV---QAEDEAQEAVVVDKRGVPIVT 177
Db 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
160 keifqkfldtknwnedrekfvgpkydmlqmdyatntqdeetkkeslkspkpes 219

Qy 178 KRVQPCSLDPATQKLTITNIFSKENFKNTMALMDLDVKKWPLGKLSKQIARGFEALE 237

```

```

Db 220 q-----ldlrveliklicnqvameemmemkntkklpklvtvaqikagyqskkie 273
Qy 238 EALKGFTDGGQSLBSLHFFYVIPHNFHGHSGOPPPINSPELLQAKMDLLVLADLELAQA 297
Db 274 dcira-gqhgrralmeacnefytriphdfglrtpplirtqksekilqillealgieiaik 332
Qy 298 LQAVSEQEKTVVEVP-HPLDRDYQLLKQQLDLSGAPEYKVIQTVLEOTGS---NHRCP 353
Db 333 l-----vktelqspehpldqhyrnlhcalrpldhesyefkvisqylqsthapthsdym 386
Qy 354 TLQHIWKVNOGEEDRFOAHSKLNKRLHWHGCTNMAVAAILTSGLRIMPH-----SGGRV 409
Db 387 tllldifevekdekeaf--edlnhmlwhgsmnvwgilshgrlrappeapitgymf 444
Qy 410 KGKIYFASENSKSAGYVIGMKCGAHHVGMFLGEVALGREHHINTDPSL----- 459
Db 445 gkglyfadmsksanycfasr--lknrtgillllsevalgcnelleanpkaeqllqgkhs 502
Qy 460 ----KSPPPGDSVTARGTEP-DPTQDTELELDGQVVVPOGQVPVCPPEFSSTFSOSE 514
Db 503 kglgknapssahfvtlmgstvpplgpasdt-----giinpdy-----tlnyne 545
Qy 515 YLIYQESQCLRYLLEV 531
Db 546 yivynpqnqrmryllkv 562

RESULT 12
AAB11480
ID AAB11480 standard; Protein; 570 AA.
XX
AC AAB11480;
XX
DT 02-MAR-2001 (first entry)
XX
DE Human brain poly-ADP-ribose-polymerase protein.
XX
KW Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW 2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW diabetes mellitus.
XX
OS Homo sapiens.
XX
PN DE19921567-A1.
XX
PD 16-NOV-2000.
XX
PF 11-MAY-1999; 99DE-1021567.
XX
PR 11-MAY-1999; 99DE-1021567.
XX
PA (BADI ) BASF AG.
XX
PI Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
DR WPI; 2001-032983/05.
XX
PD N-PSDB; AAC82090.
XX
PT Drugs for inhibiting PARP or especially homologous enzymes comprising
PT 4-substituted phthalazinone derivatives, useful e.g. for treating
PT neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS Example A; Page 12-13; 14pp; German.
XX
CC This invention describes novel 4-substituted 2H-phthalazin-1-one
CC derivatives (I) which are used for the treatment or prophylaxis of
CC diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP);

```





Db 5 gkankdrtdedkq-----esvkalllkgk-apvdpec--takvgkahvcegnvdyd 54  
 QY 68 CTLNQNIENNKKFYIIQLLQ-DSNRFFTCNWRGVRGEVQ-SKINHFTRLEDAKKDF 125  
 Db 55 vmlnqcnlgnnnkyyllqlladdagrnsvwmrvrgvmgqgshsvacsgnlkakeif 114  
 QY 126 EKRFREKTKNNWAERDHFVSHPGKYFLIEV---QAEDEAQAQEAUVKDRGPVRTVTKRQV 181  
 Db 115 qkkfldtknnwedrekfekvpgkydmlqmdyatntqdeetkkeslspkpesq--- 171  
 QY 182 PCSLDPATOKLITNIFSKEMFKNTMALMDLVKKMPLGKLSKQOIARGFEALAEALK 241  
 Db 172 ---ldlrvgelklkicnvqameemmmemkntkpkaplgltvaqikagyqslkkiedcir 228  
 QY 242 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPPTNSPELLOAKKMDLLVLADIELAQALQAV 301  
 Db 229 a-gqhgralmeacnefytriphdfgrtpplirtqkelskqllealqdieiafkI--- 284  
 QY 302 SEQEKTVEEVP-HPLDRDYQLLKCQLQLDGAPEYKVIQTYLEQTS---NHRCPQLQH 357  
 Db 285 ---vktelqspelpdqhyrnlhcalrpidhesyefkvisqylqstchapthsdymtlld 341  
 QY 358 INKVNQEGEDRFQAHSKLGNRKLLWHGNTNMAVVAAILTSGLRIMPH-----SGGRVGKG 413  
 Db 342 lfevekdgekeaf--edlhnrmllwhgsrsmnwgvlshglriappeapitgymfkgi 399  
 QY 414 YPASENKSAGYVIGMKCAHHVGYMFLGEVALGREHHINTDNPSL----- 459  
 Db 400 Yfadmssksanycfasr--lkntglllsevalgqcneleapkaegllqgkhtskglg 457  
 QY 460 KSPPPGFDSVIARGHTEP-DPTQDTELELDGQVQVVPQGPVPCPEFSSTFSQSEYLIY 518  
 Db 458 knapssahfvtngstvpqlpasdt-----gilnpdgy-----tlnyneiyiv 500  
 QY 519 QESQCRRLYLEV 531  
 Db 501 npnqvrmryllkv 513  
 RESULT 14  
 AAY68835  
 ID AAY68835 standard; Protein; 637 AA.  
 AC AAY68835;  
 XX  
 XX The poly(ADP-ribose) polymerase NAP protein of Arabidopsis.  
 DT 16-MAY-2000 (first entry)  
 XX  
 DE NAP: poly(ADP-ribose) polymerase; PARP: poly(ADP-ribose) transferase;  
 KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;  
 KW pest; drought; heat; fungi; nematode; seed-shatter.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 FH Key Location/Qualifiers  
 FT Misc-difference 1..138 /note= "these residues are specifically claimed in  
 FT claim 18"  
 XX  
 XX WO200004173-A1.  
 PN 27-JAN-2000.  
 PD 12-JUL-1999; 99WO-EPO4940.  
 XX 17-JUL-1998; 98US-0118276.  
 PR (PLBZ ) PLANT GENETIC SYSTEMS NV.  
 XX  
 PA Babyichuk E, Kushnir S, De Block M;  
 PI WPI; 2000-182436/16.  
 XX  
 XX

DR N-PSDB; AA260617.  
 XX  
 PT Modulating cell death, growth and stress resistance in eukaryotes,  
 PT specifically plants, used, e.g. to impart fungus or nematode resistance  
 XX  
 PS Claim 18; Page 99-101; 126pp; English.  
 XX  
 CC The present sequence represents a NAP protein. This protein is a  
 CC poly(ADP-ribose) polymerase (PARP) protein (also known as  
 CC cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide  
 CC sequences can be used for modulation of programmed cell death in  
 CC eukaryotic cells. The method is used, specifically in plants, to induce,  
 CC or protect against, programmed cell death, depending on the extent to  
 CC which PARP activity is reduced. Reducing expression of endogenous NAP  
 CC class PARP only is also used to modulate programmed cell death, to  
 CC increase growth rate and to produce plant cells that are more tolerant  
 CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,  
 CC etc., or during transformation). Particular applications are generation  
 CC of plants that are resistant to fungi or nematodes; are male or female  
 CC sterile; or have better seed-shatter properties. The methods are also  
 CC used to improve growth of transformed plant cells (and derived calli or  
 CC complete plants).  
 XX  
 SQ Sequence 637 AA;  
 Query Match 24.4%; Score 688.5; DB 21; Length 637;  
 Best Local Similarity 36.2%; Pred. No. 7e-54;  
 Matches 174; Conservative 89; Mismatches 185; Indels 33; Gaps 16;  
 QY 66 YNCTLNQNIENNKKFYIIQLLQ-DSNRFFTCNWRGVRGEVQSKIN-HFTRLEDAKK 123  
 Db 169 Ydallmqtnvdrnnkfvlqvlesdskktymvytrwgrvrgvkgskldgpydwdrate 228  
 QY 124 DFEKFKREKTKNNWAERDHFVSHPGKYTLIEVQAEQAQEAUVKDRGPVRTVTKRQV- 182  
 Db 229 iftnkfnkdktnywsdrkefiphkysytwlemdygkeendspvnnnd---lpssevkpe 285  
 QY 183 -CSLDPATOKLITNIFSKEMFKNTMALMDLVKKMPLGKLSKQOIARGFEALAEALK 241  
 Db 286 qskldtrvakfislcnvsmmaqhmeigynanklplgkiskstiskgyevlkrisevi- 344  
 QY 242 GPTDGGQSLEELSSHFYTVIPHNFGHSQPPPTNSPELLOAKKMDLLVLADIELAQALQ 299  
 Db 345 -drydrtrieelsgefytrviphdfgrtkmsqfvidtpqklkqkienvealgeielatkl 403  
 QY 300 AVSEQEKTVEEVPHPDLDRDYQLLKCQLQLDGAPEYKVIQTYLEQT-GSNHRCPTIQ-- 356  
 Db 404 svdpqglq-----ddplyhyqqlncgltpvgndseefsmvanymenthakthsgytveia 458  
 QY 357 HIWKVNOEGEEDRFQAHSKLGNRKLLWHGNTNMAVVAAILTSGLRIMPH-----SGGRVGKG 412  
 Db 459 qifrasraveadrfgqfsssknrmlwhgsrlltnwagilsggriappeapvgtgymfkg 518  
 QY 413 IYPASENSKSAGYVIGMKCAHHVGYMFLGEVALGREHHINTDNPSLKSPPPGFDSVIAR 472  
 Db 519 vyfadmfksanycyca-ntgand-gvlllcevalgdmnelllysdynadnlpdgklstkgv 576  
 QY 473 GHTEPDPTDTELELDGQVQVVPQGPV--PCPEFSSTFSQSEYLIYQSCRLRYLLE 530  
 Db 577 gktapnpseaqtle-dg--vvvplgkpvsc---skgmlllyneyivynveqikmryviq 630  
 QY 531 V 531  
 Db 631 v 631  
 RESULT 15  
 AAB93513  
 ID AAB93513 standard; Protein; 531 AA.  
 XX



**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**



Db 480 PWGAEVKAEPEVVPVAPRGKSGAALSKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS 537  
Qy 53 PLSSNPGTQVYED----YNCITNOTNIENNKNFYIIQLQD--SNRFFTCWNRWGRVGE 106  
Db 538 GLEHS--AHVLEKGGKVFSAITGLVDIVKGTNSYKQLLEDDEKENRYW-IFRSWGRVGT 594  
Qy 107 V-GOSKINHTRLEDAKDFEKKFREKTNNWAERDHFVSHPGKYTLIEVO-AEDEAQA 164  
Db 595 VIGSNKLEQMPSEKDAEQFMKLEETGNNAHWSK-NFTKPKFYPLEIDYGOE--EA 651  
Qy 165 VVKVDRGPVTRVTRVOPCSLDPATQKLTITNIFSKFMKNTMALMDLDVKKMPLGKLSKQ 224  
Db 652 VKLTUNP-GTKSKLPK-----VQDLIKMIFDVESMKKAMVEYEDLQKMPGLKLSKR 704  
Qy 225 QIARGFEALEALKAGPTDGGOSLELSHFYTVIPNFGHSGQPPINSPELLQAKD 284  
Db 705 QIQAAYSILSEVQAVSQGSDSOL-DLSNRFYTLPHDFGKMKPPLNNADSVQAKVE 763  
Qy 285 MLLVLADIELAQAQVSEQKTVVEVPHPLDRDYQLLKCOLQLDGAPEYKVIQTYLE 344  
Db 764 MLDNLLDIEVAYSLRGSDSDSKD-----PIDVNYEKLKTDIKVVDRODSEAEIIRKYK 819  
Qy 345 QT-GSNHRCPTLQ--HIWKVNOGEEDRFQAHSKLGNKRLHGHGTNNMAVVAAILTSLRI 401  
Db 820 NTHATTHAYDLEVIDIFKIEREGECORYKPFKQHNRLHGHGSRRTNFAGILSQGLRI 879  
Qy 402 MPH----SGGRVKGKGYFASSENSKAGYVIGMKGCAHH-----VGYMFLGEVALGREHH 451  
Db 880 APPAPVTGYMFGKGYFADVMVSKSANY-----YHTSQGDPGIGLILGEVALGNWYE 931  
Qy 452 INTDNPSLKSPPGFDSVIARGHTEPDPTQDTELELDGQVVPVQGPVPCPEFSSTFS 511  
Db 932 LKHAS-HISRLPKGKHSVGLGKTTPDPS--ANISLDG--VDVPLGTGSSGVIDTSL- 985  
Qy 512 QSEYLIQESQCRRLYLEV 531  
Db 986 YNEYIVYDIAQVNLKYLKL 1005

RESULT 2  
US-08-860-886-2  
; Sequence 2, Application US/08860886  
; Patent No. 6335009  
; GENERAL INFORMATION:  
; APPLICANT: Burkle, Alexander  
; APPLICANT: Zur Hausen, Harald  
; APPLICANT: Jan-Helner, Kupper  
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE  
; TITLE OF INVENTION: IN GENE THERAPY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds, LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSQ for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/860,886  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A.  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 8484-0028-999

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; US-08-860-886-2

Query Match 23.6%; Score 666.5; DB 4; Length 1013;  
Best Local Similarity 33.4%; Pred. No. 3.9e-59;  
Matches 186; Conservative 101; Mismatches 207; Indels 63; Gaps 24;

Qy 7 PW-----VQTEGPEKKKGRQAGR-----EDPFRSTAEALKAIPAEKRIIRVDPTC 52  
Db 479 PWGAEVKAEPEVVPVAPRGKSGAALSKSKGQVKEEGINKSEKRMKL--TLKGGAAVDPDS 536  
Qy 53 PLSSNPGTQVYED----YNCITNOTNIENNKNFYIIQLQD--SNRFFTCWNRWGRVGE 106  
Db 537 GLEHS--AHVLEKGGKVFSAITGLVDIVKGTNSYKQLLEDDEKENRYW-IFRSWGRVGT 593  
Qy 107 V-GOSKINHTRLEDAKDFEKKFREKTNNWAERDHFVSHPGKYTLIEVO-AEDEAQA 164  
Db 594 VIGSNKLEQMPSEKDAEQFMKLEETGNNAHWSK-NFTKPKFYPLEIDYGOE--EA 650  
Qy 165 VVKVDRGPVTRVTRVOPCSLDPATQKLTITNIFSKFMKNTMALMDLDVKKMPLGKLSKQ 224  
Db 651 VKLTUNP-GTKSKLPK-----VQDLIKMIFDVESMKKAMVEYEDLQKMPGLKLSKR 703  
Qy 225 QIARGFEALEALKAGPTDGGOSLELSHFYTVIPNFGHSGQPPINSPELLQAKD 284  
Db 704 QIQAAYSILSEVQAVSQGSDSOL-DLSNRFYTLPHDFGKMKPPLNNADSVQAKVE 762  
Qy 285 MLLVLADIELAQAQVSEQKTVVEVPHPLDRDYQLLKCOLQLDGAPEYKVIQTYLE 344  
Db 763 MLDNLLDIEVAYSLRGSDSDSKD-----PIDVNYEKLKTDIKVVDRODSEAEIIRKYK 818  
Qy 345 QT-GSNHRCPTLQ--HIWKVNOGEEDRFQAHSKLGNKRLHGHGTNNMAVVAAILTSLRI 401  
Db 819 NTHATTHAYDLEVIDIFKIEREGECORYKPFKQHNRLHGHGSRRTNFAGILSQGLRI 878  
Qy 402 MPH----SGGRVKGKGYFASSENSKAGYVIGMKGCAHH-----VGYMFLGEVALGREHHINT 454  
Db 879 APPAPVTGYMFGKGYFADVMVSKSANY-----CHTSQSDPIGLILGEVALGNMYELKH 933  
Qy 455 DNGSLKSPPGFDSVIARGHTEPDPTQDTELELDGQVVPVQGPVPCPEFSSTFSQSE 514  
Db 934 AS-HISRLPKGKHSVGLGKTTPDPS--ANISLDG--VDVPLGTGI--SSGVNDTSLLYNE 987  
Qy 515 YLIQESQCRRLYLEV 531  
Db 988 YIYDIAQVNLKYLKL 1004

RESULT 3  
US-09-196-387-2  
; Sequence 2, Application US/09196387  
; Patent No. 6277613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor

```

; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/196,387
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/095,225
; FILING DATE: June 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-230 CIP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1327 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; US-09-196-387-2

```

Query Match 4.9%; Score 138.5; DB 4; Length 1327;  
 Best Local Similarity 24.4%; Pred. No. 0.00011;  
 Matches 73; Conservative 41; Mismatches 102; Indels 83; Gaps 16;

```

Qy 277 ELLQAKDMLVLADI---ELAQ-ALQAVSEQEKTVEVPHPL-----DRDYQLLKQLQ 327
Db 1045 DIFETEQLDLVADMGHEELKEIGINAYGHRHKLIRGVERLLGGQOQTNPYLTFHCYNQ 1104

Qy 328 ---LIDSGAP---EYKVIQTLYEQTGSHRCP-----TLQHIWKVNOEGEEDRF- 370
Db 1105 GTILLDL-APEDKEYQSVSEEMQSTIREHRDGGNAGGIFNRYNIRQKVYVKKLRERFC 1163

Qy 371 -----QAHSKLGKRLKLLHGTNNMVAAILTSGL-RIMPHSGRGVKGKIYFASENSKS 422
Db 1164 HROKEVSEENHNHNERMLPHGS--PFINAIHKGFDERHAYIGMFCAGIYFAENSSKS 1221

Qy 423 AGYVIGM---KCGAH-----HVGVMFLGEVALGREHINTDNPSLKSPFGDSVI 470
Db 1222 NOYVYIGGGTGCPTHKDRSCYICHRQLFC-RYTLCKS-FLQFSTMKMAHAPGHHSVI 1279

Qy 471 ARGHTPEPTQDTELELDGQVVPQGPVPCPEFSSTFSQSEVLIYVOESQCLRLLYL 529
Db 1280 -----GRP-----SVNGLAYAEVYIRGEQAYPEYLI 1306

```

RESULT 4  
 US-08-923-992A-2  
 ; Sequence 2, Application US/08923992A  
 ; Patent No. 6280738  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tai, Joseph Y.  
 ; APPLICANT: Blake, Milan S.  
 ; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
 ; TITLE OF INVENTION: Streptococcal Beta Antigens  
 ; NUMBER OF SEQUENCES: 34  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
 ; STREET: 1100 New York Avenue, N.W., Suite 600

```

; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-923-992A-2

```

Query Match 4.0%; Score 113.5; DB 4; Length 1164;  
 Best Local Similarity 20.2%; Pred. No. 0.034;  
 Matches 107; Conservative 72; Mismatches 205; Indels 147; Gaps 25;

```

Qy 21 QAGREEDPFRSTAEALKAIPAERIRVDPTCPPLSSNPGTOVEDYNTCLNQTIE-NNN 79
Db 309 QATQVKNOFLENAOKLEIQ-----PLIKETNVKLYKAMESLSQVKEKELKHN 356

Qy 80 NKFYIOLLDSNRRFTCW-----NRWGRVGEVGOSKINHFRLEDAKKDFEKKFREK--- 132
Db 357 SEANLEDLVAKSKEIVREYEGKLNQSNLPELKQLEEAHKSRLQVVEDFRKFKTSQV 416

Qy 133 -----TKNNAERDHFVSHPKYTLIEVQAEDEAQAQVAVKVDGPPV---RVT 177
Db 417 TPKKRVKRDLAANENNOOKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDSDLL 473

Qy 178 KRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQIARGFEALEALE 237
Db 474 TKYNPSVSDRISTNVKTN-----TDNHKTAETIKNL---KLNESQTV----- 513

Qy 238 EALKGPTDGGOSLBELSSHFTVYIPHNFGHSQPPPIPSPELLQAKKMLLVADIQLAQA 297
Db 514 -TLKAKDSDGNVVEKT---FTITVQKKEEQVP--KTPE---QKD----- 549

Qy 298 LOAVSEQEKTVEEYP-HPLDRDYQLLKQLQ-LLDGSAPEYKVIQTYLEQTSNHRCP 355
Db 550 -----SKTEEYPOEPKSN-----KNQLELTKSAQOELEKLEKAIKL----- 589

Qy 356 QHINKVNOEGEEDRFQAHSKLGNKRLKLLHGTNNMVAAILTSGLRIMPHSGRGVKGKIYF 415
Db 590 -----MEQPEIPSNPEYGIQKSWESQKEPIQEI-TSFKKIIGDSSSKYTYEYHF 639

Qy 416 ASENSKSAGYVIGMKCGAHVGYMFLGEVALGREHHIN--TDNPSLKSPFGFDSVIAR- 472
Db 640 NKYSDFMNYQL-----HAQMEMLTRKV---QYNNKYPDNAEIKKI---FESDMKRT 686

Qy 473 -----GHTPEP-----PTQDTELELDGQVVPQGPVPCPEFS 506
Db 687 KEDNYGSLNDALKGFEKYEFLTPFNKIKQIVDDLDKKVQDQAPAPIPENS 737

```

```

Query Match      4.08; Score 112.5; DB 4; Length 717;
Best Local Similarity 19.3%; Pred. No. 0.019;
Matches 102; Conservative 81; Mismatches 229; Indels 117; Gaps 21;

QY 13 GPEKKGRQAGREEDP----FRSTAEALKAIPAEKRIIRVDPTCPLLSSNPQTQVVEDYNC 68
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 87 GGERTRRESQESDPEDDDVKKPALQSSVATSKERTFRD-----LIQDN- 133

QY 69 TLNQTNTENNKKYII-----QLQDSNRFFTCWNRWGRVGEVQSKINHFTRLEDAKK 123
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 134 -MDEKGQRNRRIFGLMGTLOKQKEST-----VATEROKRQEIEQ-KLE--VQAEERK 186

QY 124 DFEKKFKETKNNWAERDHFVSHPGKYTLIEVQAE-DEAQAVVYKVDPRGVPVTKR--- 179
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 187 QVENERRELFERRAKQTELRLLEQKVELAQLEWNEHNAKIY-----IRTKPKHLF 242

QY 180 VOPCSLPATOKLTLNIFSK--EMFKNTMALDMLDVKMPLG-----KLSKQOITARGFE 231
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 243 YIPGMCPTAKLIEESORKNNALFEGRRIEFAQINKWEARPRRQSMKEKEHQVVRNEE 302

QY 232 ALEALEALKGPTDGGOSLEELSHTYVIPHNFGHSQPPPIINSPELLEQAOKDMLLVLAAD 291
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 303 OKABOEGKYAORE--RELEETGNHNDVETEEAGEER-----EKETAIVHSD 348

```

Query Match	4.08;	Score 112.5;	DB 4;	Length 1098;
Best Local Similarity	20.08;	Pred. No. 0.039;		
Matches 107;	Conservative	75;	Mismatches 199;	Indels 153; Gaps
Qy	21	QAGREDPFRSTAEALKAIPA	EKRIIRVDP	TCPLSSNPGTOYEDYNCTLNQTNIE-NNN 79
		: :		: : : :    : :
Db	272	QATVQKQFLENAOKLEIQ	-----	PLIKETNVKLYKAMSELSQVSEKELKHN 319
Qy	80	NKEFYIOLLQDSNFFTCW	-----NRW	RGVGVGQSKNHFTLEDAKDKFEEKKPREK --- 132
		: : : :		: : : :    : :
Db	320	SEANLEDLVAKSKETIVREYEGKLN	OSKNILPEL	KOLREPAHSHKLVQVDEDFRKKFTSEOV 379



```

Qy 133 -----TKNWAERDHFVSHPGYKTYLLEVOAQEDAQEAUVVKVDGPGV---RTVT 177
      :| | : : || | | | : | | | | | :
Db 380 TPXKRLKRDLAAENNOOKTELTVS-PENITVYE--GEDVKFTVTAKSDSKTLLDFSDLL 436
      :| | : : || | | | : | | | | | :
Qy 178 KRVPQCSLDPATOKLTINIFSEKMFNTMALMDLDVKKMPLGLKLSKQIARTGFEEALE 237
      :| | : : || | | | : | | | | | :
Db 437 TKYNPSVSDRISNYKTN-----TDNHKIABITKNL---KLNSQTV----- 476
      :| | : : || | | | : | | | | | :
Qy 238 EALKGPPTGGOSLEESSHYPYTVIPHNFGHSGPPPIINSPELLQAOKMLLVADIELAQA 297
      || | : : || | | | : | | | | | :
Db 477 -TLKAKDDSGNVVEKT-----FTTVQKKEQVP--KTPE----QKD----- 512
      :| | : : || | | | : | | | | | :
Qy 298 LQAVSEOEKTVVEVP-HPLDRDYQLLKCOL-LLDSGAPYKYVIQYYLBOTGSNHRCPPL 355
      :| | : : || | | | : | | | | | :
Db 513 -----SKTEEKVPQEPKSND-----KNOQLIKSAQQELEKKAIKEL----- 552
      :| | : : || | | | : | | | | | :
Qy 356 QHIWKVNOGEEDRFQAHSLGNRKLRLLWHTGTNNMAVVAAILTSGLRIMPHSGRGVKGIVF 415
      :| | : : || | | | : | | | | | :
Db 553 -----NEQPPEIPSNPEYGIQKSIWESKEPIQEAI-TSEFKLIIGDSSKYYTEHYF 602
      :| | : : || | | | : | | | | | :
Qy 416 ASENSKAGYVIGMKCAHHVGMYMTLGEV-ALGRE--HHIN--TDNPSLKSPPPGPFDSVI 470
      :| | : : || | | | : | | | | | :
Db 603 -----NKYSHPMNQLHAQEMLETRKVYQVMKNKPYDNAEIKKI---FESDM 646
      :| | : : || | | | : | | | | | :
Qy 471 AR-----GHTEPD-----PTQDTELELDGQVVVYPOGQPVPCPEFS 506
      :| | : : || | | | : | | | | | :
Db 647 KRTKEDNYGLENDALKGYPEKLYFLTPFNKIKOIVDDFKKVEDQDPAPIPENS 700
      :| | : : || | | | : | | | | | :

RESULT 7
US-08-923-992A-6
; Sequence 6, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IqA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-6

```



QY 363 QGSEEDRFQAH-SKLGNRKLLW-----HGTNMAVVAAILTSLGRIMPHSGGRV-GK 411  
Db 1267 ----KEAMKRKHIDALLERKVRPSKPHRTN----AFIVESGTSIDPKTGKEIRCK 1314

RESULT 10

US-09-349-546-1  
; Sequence 1, Application US/09349546  
; Patent No. 6093569  
; GENERAL INFORMATION:  
; APPLICANT: Olszewski, N.  
; APPLICANT: Tzafir, I.  
; APPLICANT: Somers, D.A.  
; APPLICANT: Lockhart, B.  
; APPLICANT: Torbert, K.  
; TITLE OF INVENTION: Sugarcane bacilliform virus promoter  
; FILE REFERENCE: 600.369US2  
; CURRENT APPLICATION NUMBER: US/09/349,546  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: US 08/694,869  
; EARLIER FILING DATE: 1996-08-09  
; EARLIER APPLICATION NUMBER: PCT/IB97/01338  
; EARLIER FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1871  
; TYPE: PRT  
; ORGANISM: sugarcane bacilliform virus  
US-09-349-546-1  
  
Query Match 3.8%; Score 108.5; DB 3; Length 1871;  
Best Local Similarity 21.0%; Pred. No. 0.25;  
Matches 100; Conservative 64; Mismatches 135; Indels 177; Gaps 25;  
  
QY 33 AEALKAIPAERIRVDPTCPPLSSNPGTQVYEDYNTLNOTNIENNKNFYIQLLODSN 92  
Db 919 AEAVK--PPEKK-----SNYELLAKQLLIENSKLMEKEILIEELN 957  
  
QY 93 RFFTCNWRNGR-----VGEVGOSKINHFTREDAKDFEKKREKTKNNWAERDH-----FV 144  
Db 958 KEIKARQETKKGKELYIEASTEVENETWKSRAELFEALYNEEVKNKASTSSVTEGM 1017  
  
QY 139 ----ERDHFVSHGPKYTLLEVQA-----EDEAQEAUV-----KVDRGPFVTVTKR 179  
Db 1018 YQVQIDHL-----RKELREVEATLEVKNVYEESEEAEEVMMASAVKDEMYRFPVIEVPE 1072  
  
QY 180 VQPCSLDP-----ATQKLTINIFSKEMF-----KNTMALMDLDVK----- 214  
Db 1073 VQKVQLTALLDGTATRSCTNQVFEIEKFLQPTKFKVKGHGVNSVTYKLDQVQKDGAKLWAG 1132  
  
QY 215 ----KMPL---SKQQTARGFEALEALEALKGPTDGGQSLSEELSSHFTYVIPHN 264  
Db 1133 ENWFRLPITYVGPMTYMGKRTQMLGCFNQSLA-----GGVRLGRTVTFYKYI--- 1181  
  
QY 265 FGHSGPPPTNSPELLOAKKMDLLVLADIELAQAQVSEQ-----KTVVEVPH 313  
Db 1182 -----ASIKANEYLQAEAEETLV-----ATSEQEFINRFSMKNKRLLEEMK- 1223  
  
QY 314 PLDRDY-----QLLKQQLLDGSAPEYKVIQTYLEOTGSHRCPYTLQHIWKVN 362  
Db 1224 --EQGYMGEDTLAHNKNQICKIELRN---PDLLI-----KDKPQTLNIOK--- 1266  
  
QY 363 QGSEEDRFQAH-SKLGNRKLLW-----HGTNMAVVAAILTSLGRIMPHSGGRV-GK 411  
Db 1267 ----KEAMKRKHIDALLERKVRPSKPHRTN----AFIVESGTSIDPKTGKEIRCK 1314

RESULT 11

US-09-172-422-1  
; Sequence 1, Application US/09172422A  
; Patent No. 6300485

; GENERAL INFORMATION:  
; APPLICANT: Adams, Arwen E.  
; APPLICANT: Chiu, Choi Ying  
; APPLICANT: Duhl, David  
; APPLICANT: Gorman, Susan W.  
; APPLICANT: Leng, Song  
; APPLICANT: Sheffield, Val  
; APPLICANT: Welch, Juliet  
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED  
; CHANNEL-15 (CNGC-15) POLYNUCLEOTIDES, POLYPEPTIDES,  
; TITLE OF INVENTION: COMPOSITIONS, METHODS, AND USES THEREOF  
; FILE REFERENCE: 200130.442  
; CURRENT APPLICATION NUMBER: US/09/172,422A  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2548  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-172-422-1  
  
Query Match 3.8%; Score 108.5; DB 4; Length 2548;  
Best Local Similarity 21.0%; Pred. No. 0.42;  
Matches 88; Conservative 63; Mismatches 126; Indels 143; Gaps 25;  
  
QY 34 EALKAIPAERIRVDPTCPPLSSNPGTQVYEDYNTLNOTNIENNKNFYIQLLODSN 92  
Db 757 EILORCKEEKYSITRKNPRTPLDQG-----MNALNEKNQHDFTDI----- 798  
  
QY 93 RFFTCNWRNGRVGEVGOSKINHFTREDAKDFEKKREKTKNNWAERDH-----FV 144  
Db 799 ----AWN--GRTG--TRQSLSSGTSLLDK-----DGIFANSTSSKLLERAHGILTRKNPK 847  
  
QY 145 SHPG--KYTLIEVQA-----EDEAQEAUVKVDGRPVRTVTKRVQPCSLDPATQ--- 190  
Db 848 SKPALPKH--LLEVNSLKHLLTRTLQDRIKSLHLH-----KKKKPPSISAQFQASL 898  
  
QY 191 -KLITNIFSKEMFKNTMALMDLVKMKPLGKLSKQOIRGFEALEALE--EALKGPTDGG 247  
Db 899 SKIMETLGAQEPY--FVKCIRSNAEKPL--RFSDDLVLRLQRLYTGMLFTVQIRQSGYSK 955  
  
QY 248 OSLEELSSHFTYVIPHNFCHGSOPPPIN-----SPELLQAKKMDLLVLADIELAQA 297  
Db 956 YSQDFVSHFVLLPRNI---IPSKFNIDQPFKKNLNPNDYQVGTWTF----- 1002  
  
QY 298 LOAVSEQEKY--VEEVPH-----LDRYQQLKQQLLDGSAPEYKVIQTYLEQTS 348  
Db 1003 ---LXEQERQHLQDLLHQEVLRIILQWFRVLLCRQHFL-----HLRQAS- 1046  
  
QY 349 NHRCPYTLQHIWK--VNQEG-----EEDRFOAHSLKGNRKLWHTGTMVAAILTSGLR 400  
Db 1047 ----VTIQRFWRNYLNQKQVRDAVQKDAFV-----MASAALLQASWR 1086  
  
RESULT 12  
US-08-923-992A-10  
; Sequence 10, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005



Qy	16	KKGRQAGR--EEDPFRSTAEALKAIPAERKRIIRVDPCTPLSSNPGTQVYEDNCTNQ	73
Db	43	KKAEADQKKVEDDKQTEEKARKEASQKLNDAVLV-----QNAKEYREVQNOR	95
Qy	74	NIENNNKFYIQLQDSNRFFCTCNRWRGVEGVQSKINH-TRLEDACK--DFEKKF	129
Db	96	SKYSDAEY-----QKLTVEDSKIEKAREQDLQNK	129
Qy	130	REXTKNNAERDHFVSHPGKYTLIEVQAE-EAQEVVYKVDRGVPRTVKRVQPCSLDPA	188
Db	130	NE-----VRVVVPEPNALAEETKKAEEAKAEKVAK-----RKDYDA	167
Qy	189	TOKLTNIFSKEFMKFNPMALMDLVKKMPGLKLS-----KQIARGFEALAEALKEALK	242

db 142 PGMOOYM-ENVRVTDGOSIVAGVSLATVNTDDNAP---

**QY** 58 PGTQVYEDYN-----CTLQNTNIENNKNFYIQLLDQSDNRFFTCWNRGRVGE 106  
||| | : ||| : ||| | | |  
**Dp** 142 PGWOQYM-FNVRDGSGLVAGSLATVINDDNAP---IIQ-----NFEPG-----RVPE 186



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:11 ; Search time 36.19 Seconds  
(without alignments)  
570.255 Million cell updates/sec

Title: US-09-701-586b-4  
Perfect score: 2823  
Sequence: 1 MAPKPKPWQTEGPEKKGR.....EYLIIQESQCLRLYLEVHL 533

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2811	99.6	533	1 PPO3_HUMAN	Q9y6f1 homo sapien
2	738	26.1	559	1 PPO2_MOUSE	O88554 mus musculu
3	729.5	25.8	583	1 PPO2_HUMAN	Q9ugn5 homo sapien
4	688.5	24.4	637	1 PPO1_ARATH	Q11207 arabidopsais
5	677	24.0	1015	1 PPO1_BOVIN	P18493 bos taurus
6	674.5	23.9	996	1 PPO1_SARPE	Q11208 sarcophaga
7	674.5	23.9	1011	1 PPO1_CHICK	P26446 gallus gall
8	673.5	23.9	1012	1 PPO1_CRIGR	Q9r152 cricetulus
9	665.5	23.6	1013	1 PPO1_HUMAN	P09874 homo sapien
10	665.5	23.6	1013	1 PPO1_RAT	P27008 rattus norv
11	653	23.1	998	1 PPO1_XENLA	P31669 xenopus lae
12	649.5	23.0	1012	1 PPO1_MOUSE	P11103 mus musculu
13	638	22.6	994	1 PPO1_DROME	P35875 drosophilila
14	393.5	13.9	538	1 YQ04_CAEEL	Q09525 caenorhabdi
15	258	9.1	1724	1 PPOV_HUMAN	Q9ukk3 homo sapien
16	190.5	6.7	135	1 PPO1_ONCMA	Q08824 oncorhynch
17	126	4.5	1222	1 YNP3_CAEEL	Q10947 caenorhabdi
18	120	4.3	1938	1 MFS_AEQIR	P24733 aequipecten
19	115.5	4.1	1395	1 YS1_YEAST	P38904 saccharomyc
20	114	4.0	918	1 HXK1_BOVIN	P27595 bos taurus
21	113.5	4.0	1164	1 BAG_STRAG	P27951 streptococc
22	113.5	4.0	4540	1 DYHC_PARTE	Q27171 paramacium
23	109.5	3.9	446	1 GAG_OHVS	P16900 ovine lenti
24	108.5	3.8	382	1 MYB_AVIMB	P01104 avian myelo
25	107	3.8	1186	1 SMC_BACSU	P51834 bacillus su
26	106	3.8	724	1 HMMR_HUMAN	O75330 homo sapien
27	106	3.8	932	1 YA1A_SCHPO	Q09897 schizosacch
28	105.5	3.7	2025	1 TTC3_HUMAN	P53804 homo sapien
29	105	3.7	882	1 RA50_PYRFU	P58301 pyrococcus
30	105	3.7	1057	1 POL_STIVAI	Q02836 simian immu
31	104.5	3.7	641	1 MYB_CHICK	P01103 gallus gall
32	104.5	3.7	1403	1 VG22_HSV1	Q00105 ictaluriid h
33	104	3.7	1324	1 CUT3_SCHPO	P41004 schizosacch

34	104	3.7	1433	1 REST_CHICK	O42184 gallus gall
35	103	3.6	716	1 RRP2_IALE2	P26123 influenza a
36	103	3.6	1940	1 MYH3_RAT	P12847 rattus norv
37	102.5	3.6	640	1 MYB_HUMAN	P10242 homo sapien
38	102.5	3.6	704	1 VPS1_YEAST	P21576 saccharomyc
39	102.5	3.6	3210	1 CENF_HUMAN	P49454 homo sapien
40	102	3.6	716	1 RRP2_IALE1	P26122 influenza a
41	101.5	3.6	584	1 LIGA_HUMAN	P41214 homo sapien
42	101.5	3.6	612	1 EXO2_BPT5	P11109 bacterioph
43	101.5	3.6	874	1 SYA_PASMO	P57933 pasteurella
44	101.5	3.6	3075	1 LMAL_HUMAN	P25391 homo sapien
45	101	3.6	782	1 L100_ADE12	P36714 human adeno

ALIGNMENTS

RESULT 1	
ID PPO3_HUMAN	STANDARD; PRT; 533 AA.
AC Q9Y6F1; OSUG81;	
DT 16-OCT-2001 (Rel. 40, Created)	
DT 16-OCT-2001 (Rel. 40, Last sequence update)	
DE Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+)-ADP-ribo	
DE ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (PADPRT-3)	
DE (hPARP-3)	
GN ADPRTL3 OR PARP3 OR ADPRT3.	
OS Homo sapiens (Human).	
OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;	
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.	
OX NCBI_TaxID=9606;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC TISSUE=Fetal brain;	
RX MEDLINE=99263509; Pubmed=10329013;	
RA Johansson M.;	
RT "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA	
RT cloning of two novel poly(ADP-ribose) polymerase homologues.";	
RL Genomics 57:442-445(1999).	
RN [2]	
SEQUENCE OF 75-533 FROM N.A.	
TISSUE=Kidney;	
RA Ansgore W., Wirkner U., Mewes H.-W., Gassenhuber J., Wiemann S.;	
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.	
CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribose}(N)-acceptor -	
CC nicotinamide + {ADP-D-ribose}(N+1)-acceptor.	
CC -!- SUBCELLULAR LOCATION: Nuclear (Potential).	
CC -!- TISSUE SPECIFICITY: Widely expressed; the highest levels are in	
CC the kidney, skeletal muscle, liver, heart and spleen; also	
CC detected in pancreas, lung, placenta, brain, leukocytes, colon,	
CC small intestine, ovary, testis, prostate and thymus.	
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.	
CC -----	
CC This SWISS-PROT entry is copyright. It is produced through a collabora-	
CC tion between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC the European Bioinformatics Institute. There are no restrictions on its	
CC use by non-profit institutions as long as its content is in no way	
CC modified and this statement is not removed. Usage by and for commercial	
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/	
CC or send an email to license@isb-sib.ch).	
CC -----	
CC EMBL; AF083068; AAD29855.1; -;	
DR EMBL; AL050034; CAB43246.1; -;	
DR HSSP; P26446; 1A26.	
DR InterPro; IPR001290; PARP.	
DR InterPro; IPR004102; PARP_reg.	
DR Pfam; PF00644; PARP; 1.	
DR Pfam; PF02877; PARP; 1.	
DR Transferrase; Glycosyltransferase; NAD; Nuclear protein;	
KW ADP-ribosylation.	
FT DOMAIN 14 20 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).	
FT CONFLICT 80 80 K -> N (IN REF. 2).	

```
FT CONFLICT 171 171 A -> G (IN REF. 2).
FT CONFLICT 411 411 K -> E (IN REF. 2).
SQ SEQUENCE 533 AA; 60117 MW; 7C0AB98E64D1B9FD CRC64;

Query Match 99.6%; Score 2811; DB 1; Length 533;
Best Local Similarity 99.6%; Pred. No. 5.1e-189;
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAPKPKWVQTEGPEKKKGROAGREEDPFRSTAEALKAIPAEEKRIIRVDPTCPPLSSNPGT 60
D 1 MAPKPKWVQTEGPEKKKGROAGREEDPFRSTAEALKAIPAEEKRIIRVDPTCPPLSSNPGT 60
QY 61 QVYEDYNCTLNQTNENNNKFFIYLIQLQDSNRFETCWNRMGRVGEVQSKINFTLEL 120
D 61 QVYEDYNCTLNQTNENNNKFFIYLIQLQDSNRFETCWNRMGRVGEVQSKINFTLEL 120
QY 121 AKKDFEKKFKREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQAQVVKVDRGPRVTVTKRV 180
D 121 AKKDFEKKFKREKTKNNWAERDHFVSHPGKYTLIEVQAEDEAQAQVVKVDRGPRVTVTKRV 180
QY 181 QPCSLDPATQKLTINIFSKEMFKNTMALMDLVKMKPLGKLSKQOIARGFEALEEAL 240
D 181 QPCSLDPATQKLTINIFSKEMFKNTMALMDLVKMKPLGKLSKQOIARGFEALEEAL 240
QY 241 KGPDTGGQSLSELSHFYTVIPNFGHSQPPPIINPELLQAKKMLVLADELIAQALQA 300
D 241 KGPDTGGQSLSELSHFYTVIPNFGHSQPPPIINPELLQAKKMLVLADELIAQALQA 300
QY 301 VSEGEKTVVEVPHPLDRDYQLLQKQLDLSGAPYKVIQYLYEQTGNSNRHRCPTLQHTWK 360
D 301 VSEGEKTVVEVPHPLDRDYQLLQKQLDLSGAPYKVIQYLYEQTGNSNRHRCPTLQHTWK 360
QY 361 VNQGEEDRFOAHSKLGKRLKLTWGTMAVVAAILTSLRTPHSGRGVGIYFASNS 420
D 361 VNQGEEDRFOAHSKLGKRLKLTWGTMAVVAAILTSLRTPHSGRGVGIYFASNS 420
QY 421 KSAGYVIGMKGAHVGYMFLGEVALGREHINTDNFSLKSPFPFGFDSVARGHTEPDPT 480
D 421 KSAGYVIGMKGAHVGYMFLGEVALGREHINTDNFSLKSPFPFGFDSVARGHTEPDPT 480
QY 481 QDTELELDGQVVPQGPVPCPFSSSTFSQSEYLYIQESQCRRLVLEVLH 533
D 481 QDTELELDGQVVPQGPVPCPFSSSTFSQSEYLYIQESQCRRLVLEVLH 533

RESULT 2
PPO2_MOUSE STANDARD; PRT; 559 AA.
ID PPO2_MOUSE 559 AA.
AC O88554; Q99N29;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-ribose transferase-2) (Poly[ADP-ribose] synthetase-2) (PADPRT-2) (mPARP-2).
GN ADPRT12 OR PARP2 OR ADPRT2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.
RC TISSUE=Embryo;
RX MEDLINE=9292755; PubMed=10364231;
RA Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apio F., Decker P., Muller S., Hoyer T., Menissier-de Murcia J., de Murcia G.M.;
RT "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose) polymerase.";
RL J. Biol. Chem. 274:17860-17868(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=129SV;
```

```
RX MEDLINE=21179160; PubMed=11133988;
RA Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M., Niedergang C.P.;
RT "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2 (PARP-2) gene to the gene for RNase P RNA.";
RL J. Biol. Chem. 276:11092-11099(2001).
RN [3]
RP SEQUENCE OF 9-559 FROM N.A.
RC STRAIN=129/SV X C57BL/6;
RX MEDLINE=9928466; PubMed=10338144;
RA Berghammer H., Ebner M., Marksteiner R., Auer B.;
RT "ADPRT-2: a novel mammalian poly(ADP-ribose) transferase gene related to truncated PADPRT homologues in plants and Caenorhabditis elegans.";
RL FEBS Lett. 449:259-263(1999).
CC -1- FUNCTION: HAS DNA-DEPENDENT POLY[ADP-RIBOSE] POLYMERASE ACTIVITY. SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.
CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor = nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.
CC -1- SUBCELLULAR LOCATION: Nuclear.
CC -1- TISSUE SPECIFICITY: widely expressed; the highest levels were in testis followed by ovary.
CC -1- INDUCTION: By high levels of DNA-damaging agents.
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ007780; CAA07679.1; -
DR EMBL: AF191547; AAK13253.1; -
DR EMBL: AF072521; AAC25415.1; ALT_INIT.
DR HSSP: P26446; 1A26.
DR MGD: MGI:1341112; Adprt2.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR Pfam: PF00644; PARP.1.
DR Pfam: PF02877; PARP_reg.1.
DR Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein; ADP-ribosylation.
FT DNA_BIND 1 65 POTENTIAL.
FT DOMAIN 66 559 NAD-BINDING (BY SIMILARITY).
FT DOMAIN 3 9 NUCLEAR LOCALIZATION SIGNAL 1ST PART (POTENTIAL).
FT DOMAIN 33 39 NUCLEAR LOCALIZATION SIGNAL 2ND PART (POTENTIAL).
FT CONFLICT 82 82 L -> V (IN REF. 2).
FT CONFLICT 177 177 V -> I (IN REF. 2).
FT CONFLICT 486 486 R -> O (IN REF. 2).
SQ SEQUENCE 559 AA; 63396 MW; E0AEDAEE412C1445 CRC64;

Query Match 26.1%; Score 738; DB 1; Length 559;
Best Local Similarity 34.3%; Pred. No. 3.5e-44;
Matches 187; Conservative 103; Mismatches 191; Indels 64; Gaps 19;

QY 16 KKKGRQA-GREEDPFRSTAEALKAIPAEEKRIIRVDPTCPPLSSNPGTQVY-----EDYNCTL 70
D 42 QKKGPMAGGKADRTKDNRSVKTLLKKGK-APVDPECAARKLGK-AHVYCEGDDVDVYML 99
QY 71 NOTNIENNNKFFIYLIQLQDSNRFETCWNRMGRVGEVQSKINFTLELDAKDKFEK 128
D 100 NQTLQNNNNKFFIYLIQLQDQAQRNFVMMRMGRVGTGQHSVLTCSGDLNKAKEIFQK 159
QY 129 FREKTKNNWAERDHFVSHPGKYTLIEV-----QAEDEA---QEAQVVKVDRGPRVTVTKRV 181
D 160 FLDTKNNWEDRENFEKVPCKYDMLQMDYAASTQDESKTEETLKPE----- 207
QY 182 PCSLDPATQKLTINIFSKEMFKNTMALMDLVKMKPLGKLSKQOIARGFEALEEAL 241
```





```

Db 277 AGYSLKKIEDCIRA-CQHGRALMEACNEFYTRIPHDGFLRTPPLRTQKELSEKIQLE 335
QY 288 VLADIELAQAQVSOEKTVEVP-HPLDRDYQLKQCQLLDGAPKVKVITQLEQT 346
Db 336 ALGDIEIAIKL-----VKTELQSPHPLDQHYRNLCALRPLDHSYEFKYSQYLOST 389
QY 347 GS---NHRCPQLQHIKVNQEGEDRFQAHSKLGNKLLWHGTNMAVVAALTSGLRIMP 403
Db 390 HATHSDTYMTLLDLFEVKGKEAFR--EDLHNRMLLWHGSRMNVWGILSHGLRIAP 447
QY 404 H-----SGRVRGKGIYFASENSKAGYVIGMKGAHVGVMFLGEVALGREHHINTDNPSL 459
Db 448 PEAPITGYMGKGIYFADMSKSKSANTCFASR--LKNTGLLLSEVALGQCNELEANPKA 505
QY 460 -----KSPPPGFDVSIARGHTEP-DPTQDTELELDGQVVVPPQGPVPCPE 504
Db 506 EGLQKHKSHKGLKMAPSSAHEFTVLTGSLVPLGPASDT-----GILNPDGY----- 552
QY 505 FSSSTFSSSELYIQESOCRLRYLLEV 531
Db 553 ---TLNNEYIYVNPQVPMRYLLKV 575

RESULT 4
PPOL_ARATH STANDARD; PRT; 637 AA.
AC Q11207; 1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-
DE ribosyltransferase) [Poly[ADP-ribose] synthetase].
GN APP.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95269779; PubMed=7750552;
RA Lepiniec L., Babiychuk E., Kushnir S., van Montagu M., Inze D.;
RT "Characterization of an Arabidopsis thaliana cDNA homologue to animal
RT poly(ADP-ribose) polymerase.";
RL FEBS Lett. 364:103-108(1995).
CC -I- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -I- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-RIBOSYL}(N)-acceptor =
CC nicotinamide + {ADP-D-RIBOSYL}(N+1)-acceptor.
CC -I- SUBCELLULAR LOCATION: Nuclear.
CC -I- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z48243; CAAB88288.1; -
CC HSSP; P26446; 1A26.
CC InterPro; IPR001290; PARP.
CC InterPro; IPR004102; PARP_reg.
CC InterPro; IPR003034; SAP.
CC Pfam; PF00644; PARP; 1.
CC Pfam; PF02877; PARP_reg; 1.
CC Pfam; PF02037; SAP; 2.

```

```

DR SMART; SM00513; SAP; 2.
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation.
FT DNA_BIND 1 140 POTENTIAL.
FT DOMAIN 140 637 NAD-BINDING (BY SIMILARITY).
FT DOMAIN 41 62 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).
SQ SEQUENCE 637 AA; 72175 MW; 527A8F464605D127 CRC64;

Query Match 24.4%; Score 688.5; DB 1; Length 637;
Best Local Similarity 36.2%; Pred. No. 1.2e-40;
Matches 174; Conservative 89; Mismatches 185; Indels 33; Gaps 16;

QY 66 YNCTLNOTNIENNKNFYITQLIQ-DSNRFCTCNRWGRVGEVQSKIN-HFTRLEDAKK 123
Db 169 YDAILNQTWVRDNNKFFVLQVLESDSKTYVYTWGRVGVKQSKLDGPDYSDWRAIE 228
QY 124 DFEKFKREKTKNNWAPRDHFVSHPGKYTLIEVOAEDAQAVVYKVRGPTVTKRVQP- 182
Db 229 IFTNKFNDKTKNYWSDRKEFIPHKSYTWLEMDYKGEENDSPVND---IPSSSEVKPE 285
QY 183 -CSLDPATQKLTINIFSKEMFKNTMALMDLVKMKPLGKLSKQOIARGFEALEALAK 241
Db 286 QSKLDTRVAKFTSLICNVSMMAQHMMEIGYNANKPLGKISKSTISKGYEVLRKRISEVI- 344
QY 242 GPTDGGOSLEELSSHFYTVIPHNFGHSQPPP--INSPELIQAQKDMLLVLADIELAQAQ 299
Db 345 -DRYDTRLEELSGEFTYVPHDFGFGKKNQSFVIDTFQKLOKIEWEALGEIELATKLL 403
QY 300 AVSEKQKTVVEVPHPLDRDYQLKQCQLLDGAPKVKVITQLEQT-GSNHRCPTLQ-- 356
Db 404 SVDPLGLQ-----DDPLYHYQQLNGLTPVGNDSSEFMSVANTMENTHAKTHSGYTVIEA 458
QY 357 HIWKVQEGEEDRFQAHSKLGNKLLWHGTNMAVVAALTSGLRIMP-----SGGRVGK 412
Db 459 QLFPRASRAVEADRFQFSSSKNRLMLWHGSRNLTNWAGILSQGRLRAPPEAPVGYMGK 518
QY 413 IYFASENSKSAGYVIGMKGAHVGVMFLGEVALGREHHINTDNPSLKSPPGFDSVIAR 472
Db 519 VYFADFMSKSNYCXA-NTGAND-GVLLCEVALGDMNELLYSDYNADNLPCKLSTKGV 576
QY 473 GHTEPPTQDTELELDGQVVVPPQGPV--PCPEFSSSTFSSSELYIQESQCLRYLLE 530
Db 577 GKTAPNPSEAQTLE-DG--VVVPLGRPVRSRSC---SKGMLLYNEYIYVNVYEQIKMRYVIQ 630
QY 531 V 531
Db 631 V 631

RESULT 5
PPOL_BOVIN STANDARD; PRT; 1015 AA.
AC P18493; Q9TS00;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE ADP-ribosyltransferase-1) [Poly(ADP-ribose) synthetase-1].
GN ADPRT.
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovoidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90382673; PubMed=2119324;
RA Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;
RT "Cloning of a full-length cDNA encoding bovine thymus
RT poly(ADP-ribose) synthetase: evolutionarily conserved segments and
RT their potential functions.";
RL Gene 90:249-254(1990).

```

RN [2]  
 RP SEQUENCE OF 647-714 AND 838-903 FROM N.A.  
 RX MEDLINE=88151954; PubMed=2450019;  
 RA Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,  
 RA Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.;  
 RT Depression in gene expression for poly(ADP-ribose) synthetase during  
 RT the interferon-gamma-induced activation process of murine macrophage  
 RT tumor cells.\*;  
 RL Eur. J. Biochem. 171:571-575(1988).  
 CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =  
 CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.  
 CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -!- SUBUNIT: HOMODIMER (Potential).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
 CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D90073; BAA14114.1; -;  
 DR EMBL; X06986; CAA30046.1; -;  
 DR EMBL; X06987; CAA30047.1; -;  
 DR PIR; JS0428; JS0428.  
 DR PIR; S00328; S00328.  
 DR HSP; P26446; 1A26.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR001510; Znf-PARP.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF02877; PARP; 1.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF00645; Zf-PARP; 2.  
 DR ProDom; PD004675; Znf-PARP; 2.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS0172; BRCT; 1.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 DR PROSITE; PS00664; PARP\_ZN\_FINGER\_2; 2.  
 DR Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 DR ADP-ribosylation; Zinc-finger; Zinc.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DNA\_BIND 1 374 AUTOMODIFICATION DOMAIN.  
 FT DOMAIN 375 525 BRCT.  
 FT DOMAIN 386 462 NAD-BINDING.  
 FT DOMAIN 526 1015 PARP-TYPE.  
 FT ZN\_FING 20 55 PARP-TYPE.  
 FT ZN\_FING 127 164 PARP-TYPE.  
 FT DOMAIN 209 211 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
 FT DOMAIN 223 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
 FT MOD\_RES 408 414 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 414 414 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 436 436 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 446 446 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 449 449 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 457 457 ADP-RIBOSYL[N] (POTENTIAL).

FT MOD\_RES 472 472 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 485 485 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 489 489 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 492 492 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 514 514 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 515 515 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 521 521 ADP-RIBOSYL[N] (POTENTIAL).  
 SQ SEQUENCE 1015 AA; 113355 MW; 0A5FE9D9F04F5B04 CRC64;  
  
 Query Match 24.0%; Score 677; DB 1; Length 1015;  
 Best Local Similarity 33.8%; Pred. No. 1.4e-39;  
 Matches 187; Conservative 99; Mismatches 20; Indels 62; Gaps 24;  
  
 QY 1 MAPKPK----PWOTEPGPKKKKGROAGREEDPFRSTAEALKAIPAERKRIIRVDPCTPLSS 56  
 DB 494 VPKPKSGAAPSKKSGPKVEEG--TNKSEKRMKLTLLGGAA-----VDPDSGLGH 542  
  
 QY 57 NPGTOVED----YNTLTNQTNIENNKNFYIQLLODSNRF-FTCWNRWGRGVEV-GQS 110  
 DB 543 N--AHVLEKGGKVSATLGLVDIVKGTNSYKQLQLEDDKESRYWIFRSWGRVGTIGSN 600  
  
 QY 111 KINHTRLEDAKDEPKFKREKTKNNAERDHFVSHPGKYTLIEVQ-AEDEAQEAHVVD 169  
 DB 601 KLEQMPSKEDALEHFMKLYEEKTGNASHK-NFTKHPKKFYPLEIDYQDE--EAVKKLT 657  
  
 QY 170 RGPVTVTKRVQCSLDPATOKLITNIEFKEMFKMTMALMDLDVKKMPLGLKSKQIARG 229  
 DB 658 VNP-GTKSKLPRP-----VONLKMIFDVESMKAMVEYEDLQKMLGLSKRQIQAA 710  
  
 QY 230 FEALEALSEAL-KGTDGQSGSLELSHFYTVIPHNFHSGQPPPIPSPELLQAKKMDLLV 288  
 DB 711 YSILSEVQALSQSSD--SHILDSNRFYTLIPHDFGKMKPPLNNSVQAKVEMLDN 768  
  
 QY 289 LADIELAQAQVSRQETVEVPHPLDRYQLKQCLQDLSGAPKVKYQTYLEQT-G 347  
 DB 769 LLDIEVAYSLLRGSDSDSKD----PIDVNEKLTDKIVVDKDSSEAEIIRKVKNTHA 824  
  
 QY 348 SNHRCPTLQ--HIKVNQEGEEDFQAHSKGLNKLKLGHTNMAVVAAILLSGLRIMPH- 404  
 DB 825 TTHNAYDLEVVDIFKIEREGESQRYKPKPKQLHNRRLMHGSRNTTNFAGILSGGLRIAPPE 884  
  
 QY 405 ---SGRVGKGIYFASENSKSAGYVIGMCKGAH---VGYMFLGEVALGREHHINTDNPS 458  
 DB 885 APVTGYMEFGKGIYFADMYSKSANY-----CHTSQGDPIGLILLGEALGNMYELKHAR-H 938  
  
 QY 459 LKSPPPGDFSVIARGHTPEPTQDTELELDGQVVVPGQVPCPEFSSTFSQSEYLIY 518  
 DB 939 ISKLPKKGKHSVGLGKTTDPDS--ASITVDG--VEVPLGTGI-SSGVNDTCLLYNEYIVY 993  
  
 QY 519 QESQCRRLRYLLEV 531  
 DB 994 DIAQVHLKLLKL 1006  
  
 RESULT 6  
 PPOL\_SARPE STANDARD; PRT; 996 AA.  
 AC Q11208;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly (ADP-ribose) polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
 DE ribosyltransferase [Poly(ADP-ribose) synthetase].  
 OS Sarcophaga peregrina (flesh fly) (Boettcherisca peregrina).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Oestroidea; Sarcophagidae; Sarcophaga.  
 OX NCBI\_TaxID=7386;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=94170813; PubMed=8125121;  
 RA Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,

de Prati A.C., Kurata S., Natori S., Sugimura T., Esumi H.;  
"Cloning and functional expression of poly(ADP-ribose) polymerase  
from *Sarcophaga peregrina*.";  
Eur. J. Biochem. 220:607-614(1994).  
-1- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
-2- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =  
nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.  
-3- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
-4- SUBCELLULAR LOCATION: Nuclear.  
-5- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
-6- SIMILARITY: BELONGS TO THE PARP FAMILY.  
-7- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: D16482; BAA03943.1; .  
CC HSSP: P26446; 1A26.  
CC InterPro: IPR001357; BRCT.  
CC InterPro: IPR001290; PARP.  
CC InterPro: IPR004102; PARP\_reg.  
CC InterPro: IPR001510; Znf-PARP.  
CC Pfam: PF00533; BRCT; 1.  
CC Pfam: PF00644; PARP; 1.  
CC Pfam: PF02877; PARP\_reg; 1.  
CC Pfam: PF00645; Znf-PARP; 2.  
CC ProDom: PD004675; Znf-PARP; 2.  
CC SMART: SM00292; BRCT; 1.  
CC SMART: PS00172; BRCT; 1.  
CC PROSITE: PS00347; PARP\_ZN\_FINGER\_1; FALSE\_NEG.  
CC PROSITE: PS00347; PARP\_ZN\_FINGER\_2; 2.  
CC PROSITE: PS00064; PARP\_ZN\_FINGER\_2; 2.  
CC TRANSFERASE: Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
ADP-riboseylation; Zinc-finger; Zinc.  
CC DNA\_BIND 1 369 BY SIMILARITY.  
FT DOMAIN 370 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 382 456 BRCT.  
FT DOMAIN 508 996 NAD-BINDING.  
FT ZN\_FING 19 54 BY SIMILARITY.  
FT ZN\_FING 126 164 BY SIMILARITY.  
FT DOMAIN 211 214 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 232 235 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT SEQUENCE 996 AA; 113018 MW; 690DDD36E7487298 CRC64;  
-----  
Query Match 23.9%; Score 674.5; DB 1; Length 996;  
Best Local Similarity 33.6%; Pred. No. 2.1e-39;  
Matches 179; Conservative 104; Mismatches 195; Indels 55; Gaps 22;  
-----  
QY 25 EEDPFRSTAEALKAIPAEKRI-----IRVDPCTPLSSNPGTQYV-----EDYNCNLQTN 74  
DB 490 ESKSKSKSYTKSVKPSMTLKIKGLAVDPDPSGLEL--VAHVTVSRNKKYINVLGITD 547  
QY 75 IENNKNFVIQLQD--SNRFETCNWRGRVG--EVQCSKINFTREDAKDFEKKFRE 131  
DB 548 IQKNKNSFYKLQLESMDKNRFVW--FRSWGRICTIGNKLDNFSNLVDALVQFKELYLE 606  
QY 132 KTKNWAERDHFVSHPKYTLIEVQ--AEDEAQAQVAVKVRGPTVTVTKRVPQCSLDPATQ 190  
DB 607 KSGHNFENFENFKVAGRMYPIDIDYAEAS-----KIDLSAEHDIKSL-PLSV-----Q 655

QY 191 KLITNIFSKEMFKNTMALMDLVKKMPLGKLSKQIQARGFEALAEALAEALKGPTDGGQSL 250  
DB 656 DIKLMPDVSMDRTMEFDLMKPLGKLSQKIQSAYKVLTEIYELIQG---GGTNA 712  
QY 251 E--ELSSHFTVTPHNFHSGQPPPIPSPELLQAKKMLLVADIELAAQALQAVSEQKTV 308  
DB 713 KFDATNRFYTLIPHNFGTQSPPLDTEQVEQLRQMLDSLIEICAYSLQTEDSKADI 772  
QY 309 EEPHPLDRDYQLLKQQLQSLDGAPEYKVIQTYLEOT--GSNHRCPVLQ--HWKVNQEG 365  
DB 773 ----NPIDKHYEOLKTKLEPLDKNSEYIILQYKVNTHAETHKLYDLEVVDFIKVARQG 828  
QY 366 EEDRFOAHSKGLNRKLLWHGTCNMAVAAILTSLGRIMPH----SGRGVKGIFYASENSK 421  
DB 829 EARYPFKLLHNRLLWHGSRITNFAGILSHGLKTAPPEAPVTGYMFGKGIYFADMVSK 888  
QY 422 SAGYVGMKCGAHH--VGYMFLGEVALGREHHINTDNPISLSPPPGSDVSIVARHTEPDP 479  
DB 889 SANYC----CTSHNSTGLMLLSEVALGDMMECTAAKYVTKLPNDK--HSCFGRGRTMPNP 943  
QY 480 QTDELELDGQVVPQGVPCPEFSSSTFSQSEYLIYQESQCRRLAYLLEVH 532  
DB 944 SESIIRE-DG--VEIPLGKPTITNDSLKSSL--YNEFTIYDIAQVNIQYMLRMN 992  
-----  
RESULT 7  
PPOL\_CHICK STANDARD; PRT; 1011 AA.  
ID AC P26446;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 15-JUL-1998 (Rel. 36, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE POLY [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
DE ADP-ribosyltransferase-1) (Poly(ADP-ribose) synthetase-1).  
OS ADPRT.  
GN Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RS SEQUENCE FROM N.A.  
RC TISSUE=Oviduct.  
RX MEDLINE=91340148; PubMed=1840535;  
RA Ittel M.-E., Garnier J.-M., Jeltsch J.-M., Niedergang C.;  
RT "Chicken poly(ADP-ribose) Synthetase: complete deduced amino acid  
RT sequence and comparison with mammalian enzyme sequences.";  
RL Gene 102:157-164(1991).  
RN [2]  
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 659-1011.  
RX MEDLINE=96353841; PubMed=8755499;  
RA Ruf A., Mennissier de Murcia J., de Murcia G.M., Schulz G.E.;  
RT "Structure of the catalytic fragment of poly(AD-ribose) polymerase  
RT from chicken.";  
RL Proc. Natl. Acad. Sci. U.S.A. 93:7481-7485(1996).  
RN [3]  
RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 559-1011, AND REVISION TO  
RP 895.  
RX MEDLINE=98191351; PubMed=9521710;  
RA Ruf A., de Murcia G.M., Schulz G.E.;  
RT "Inhibitor and NAD+ binding to poly(ADP-ribose) polymerase as derived  
RT from crystal structures and homology modeling.";  
RL Biochemistry 37:3893-3900(1998).  
RN [4]  
RP X-RAY CRYSTALLOGRAPHY (2.25 ANGSTROMS) OF 659-1011.  
RX MEDLINE=98239716; PubMed=9571033;  
RA Ruf A., Rolli V., de Murcia G.M., Schulz G.E.;  
RT "The mechanism of the elongation and branching reaction of poly(ADP-  
RT ribose) polymerase as derived from crystal structures and  
RT mutagenesis.";  
RL J. Mol. Biol. 278:57-65(1998).  
CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR

CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose](N)-acceptor -  
 CC nicotinamide + [ADP-D-ribose](N+1)-acceptor.  
 CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -1- SUBUNIT: HOMODIMER (Potential).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
 CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X52690; CAA36917.1; --  
 CC PIR; JH0581; JH0581.  
 CC PDB; 2PAW; 27-MAY-98.  
 CC PDB; 1PAX; 13-MAY-97.  
 CC PDB; 2PAX; 27-MAY-98.  
 CC PDB; 3PAX; 27-MAY-98.  
 CC PDB; 4PAX; 27-MAY-98.  
 CC PDB; 1A26; 27-MAY-98.  
 CC InterPro; IPR001357; BRCT.  
 CC InterPro; IPR001290; PARP.  
 CC InterPro; IPR004102; PARP-reg.  
 CC InterPro; IPR001510; Znf-PARP.  
 CC Pfam; PF00533; BRCT. 1.  
 CC Pfam; PF00644; PARP. 1.  
 CC Pfam; PF02877; PARP-reg. 1.  
 CC Pfam; PF00645; 2f-PARP; 2.  
 CC ProDom; PD004675; Znf-PARP; 2.  
 CC SMART; SM00292; BRCT. 1.  
 CC PROSITE; PS01172; BRCT. 1.  
 CC PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 CC PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
 CC TRANSFERASE; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 CC ADP-ribosylation; zinc-finger; zinc; 3D-structure.  
 CC  
 CC DNA\_BIND 1 370  
 CC FT DOMAIN 371 522 AUTOMODIFICATION DOMAIN.  
 CC FT DOMAIN 382 458 BRCT.  
 CC FT DOMAIN 523 1011 NAD-BINDING.  
 CC FT ZN\_FING 21 56 PARP-TYPE.  
 CC FT ZN\_FING 125 162 PARP-TYPE.  
 CC FT DOMAIN 207 209 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
 CC FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
 CC FT MOD\_RES 403 403 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 404 404 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 410 410 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 411 411 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 432 432 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 434 434 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 441 441 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 442 442 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 453 453 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 454 454 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 468 468 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 481 481 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 485 485 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 488 488 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 509 509 ADP-RIBOSYL[N] (POTENTIAL).  
 CC FT MOD\_RES 510 510 ADP-RIBOSYL[N] (POTENTIAL).

FT MOD\_RES 517 517 ADP-RIBOSYL[N] (POTENTIAL).  
 FT CONFLICT 895 895 A -> R (IN REF. 1).  
 SQ SEQUENCE 1011 AA; 113520 MW; 261AED9383139144 CRC64;  
 Query Match 23.9%; Score 674.5; DB 1; Length 1011;  
 Best Local Similarity 33.4%; Pred. No. 2.1e-39;  
 Matches 187; Conservative 100; Mismatches 204; Indels 69; Gaps 24;  
 Qy 7 PW---VQTEGPE-----KKGRQAGREDPPRSTAEALKAIPAKRIIRVDPT 51  
 Db 477 PWGAEVTEHOEAVDVGKCSKPANKSAGKVEQPSKSEKKMLTV---KGAADVDP 533  
 Qy 52 CPLSSNPGCTOYED---YNCTLQNTNIENNKNFYIIQLQD--SNRFFCTWNWGRVG 105  
 Db 534 SGLEDS--AHVFEKGGKIFSATLGLVDIVKGTNSYKQLLEDDEDSRYVW-FRSWGRVG 590  
 Qy 106 EV-QQSKINHFTRLEDAKDFEKKFREKTKNNWAERDHFVSHGPKYTLIEVO-ADEAQE 163  
 Db 591 TVIGSNKLEQMPSEDAVEHFLNLYEKTGNSWHSK-NFTKYPKFYPLEIDYGDEEAV 649  
 Qy 164 AVVVDGRGPVTIVKRVOPCSLDPATQKLTINIFSKEMFKNTMALMDLDVKKMPLGLSK 223  
 Db 650 RKLTVSAG---TKSLAKP-----IQDLIKMIFDVESMKKAMVEFIDLQKMPGLKLSK 700  
 Qy 224 QOIARGFEALAEALKGPDTGGOSLE--ELSSHFTVVIPIHNFHGSQPPPIINSPELLQA 281  
 Db 701 RQIQSAYSILNEVQAV---SDGGSSEQLDLSNRFYTLIPHDFGKFKPILLSNLEYQA 757  
 Qy 282 KKDMLLVLADELAAQLOAVSEQKTVVEVPHPLDRDYOLLKCOLQLDSDGAPEYKVTOT 341  
 Db 758 KVMQDLNLLDIEAVYSLLRGNEDGDKD---PIDINYEKLRDIDKVVVDKDEEAKIKQ 813  
 Qy 342 YLEQT-GSNHRCPTLO--HIWKVNOGEEDRFQAHSKLGNRKLWHGNTNMAVVAAILTSG 398  
 Db 814 YVKNTHAATHNAYDLKVVVEIFRIEREGESQYKPKQLHNRQLLWHGSRNTNFAGILSQG 873  
 Qy 399 LRIMPH----SGGRVKGIGYFASENSKSAGYVIGMKC---GAHHVGYMFLGEVALGRHH 451  
 Db 874 LRIAPPEAPVTGVWFGKGIYFADWVSKSANY-----CHTSQADPIGLILLGEVALGNVYE 928  
 Qy 452 INTDNPSLKSPPGFSDSVIARGHTPEPTQTELELDGQVVPQGVQVPCPEFSSSTFS 511  
 Db 929 LKNASHITKL-PKGKHSVKGGLKTAAPTATT--TLDG--VEVPLNGIT-STGINDTCLL 982  
 Qy 512 QSEYLYQESOCRLRYLLEV 531  
 Db 983 YNEYIVYDVAQVNUKYLKLL 1002

RESULT 8  
 PPOL\_CRIGR STANDARD; PRT; 1012 AA.  
 ID PPOL\_CRIGR  
 AC Q9R152;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
 DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
 GN ADPRT.  
 OS Cricetus griseus (Chinese hamster).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 OC Cricetulus.  
 OC NCBI\_TaxID=10029;  
 [1]  
 SEQUENCE FROM N.A.  
 RX MEDLINE=21276334; PubMed=11382339;  
 RA Ganesh A., Phillips E., Thacker J., Meuth M.;  
 RT "Suppression of the radiation-sensitive phenotype of hamster irs1 and  
 RT irs2 strains selected for resistance to 3-aminobenzamide.";  
 RL Int. J. Radiat. Biol. 77:609-616(2001).  
 CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR

CC PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (BY  
 CC similarity).  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-riboseyl}(N)-acceptor =  
 CC nicotinamide + {ADP-D-riboseyl}(N+1)-acceptor.  
 CC -1- COFACTOR: ZINC. CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY  
 CC SIMILARITY).  
 CC -1- SUBUNIT: HOMODIMER (Potential).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
 CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AF168781; AAD45817.1; -;  
 DR HSSP; P26446; 1A26.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR001510; Znf-PARP.  
 DR Pfam; PF005333; BRCT; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF00645; zf-PARP; 2.  
 DR ProDom; PD004675; Znf-PARP; 2.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS0172; BRCT; 1.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
 KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 KW ADP-ribosylation; Zinc-finger; Zinc.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DNA\_BIND 1 371 BY SIMILARITY.  
 FT DOMAIN 372 522 AUTOMODIFICATION DOMAIN (BY SIMILARITY).  
 FT DOMAIN 386 475 BRCT.  
 FT DOMAIN 523 1012 NAD-BINDING (BY SIMILARITY).  
 FT ZN\_FING 20 55 PARP-TYPE.  
 FT ZN\_FING 124 161 PARP-TYPE.  
 FT DOMAIN 206 208 NUCLEAR LOCALIZATION SIGNAL 1ST PART  
 FT DOMAIN 220 225 (POTENTIAL).  
 FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART  
 FT (POTENTIAL).  
 FT MOD\_RES 406 406 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 412 412 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 434 434 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 436 436 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 443 443 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 444 444 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 455 455 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 470 470 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 483 483 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 487 487 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 490 490 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 511 511 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 512 512 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 518 518 ADP-RIBOSYL[N] (POTENTIAL).  
 SQ SEQUENCE 1012 AA; 112400 MW; 8D617C4DBF0CB0F7 CRC64;

Query Match 23.9%; Score 673.5; DB 1; Length 1012;  
 Best Local Similarity 33.8%; Pred. No. 2.5e-39;

Matches 186; Conservative 100; Mismatches 205; Indels 59; Gaps 22;  
 QY 2 APKPK---PWQTEGPEKKKGROAGREEDPFRSTAEAKAIPAERKRIIRVDTCPPLSSNP 58  
 DB 493 APGKSAAPSKKSGLYKEEG--VNKSEKRMKLTGKGAA-----VDPGSLGHS- 540  
 QY 59 GTQVYED----YNCNTQNTNIENNKNKFIYIQLQDSNRF-FTCWNRWGRVGEV-GQSKI 112  
 DB 541 -AHVLEKGGKVFATGLDIVKGTNSYYKLQLEDDKESRYWIFRSWGRGVIGSNKL 599  
 QY 113 NHTRLDAAKDFEKKFREKTKNNWAERDHFVSHPGKYTLIEVQ-AEDEAQEAQVYKVDRG 171  
 DB 600 EQMPSEKEDAVEHEMKLYEEKTGNAAWSK-NFTKYPKKFYPLEIDYGQDE--EAVKKL--- 653  
 QY 172 PRTVTKRVQPCSLDPATOKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGFE 231  
 DB 654 TVKPGTK-----SKLPKAVQELVGMIFDVESMKKALVEYIDLOKMPGLGKLSKQIQAAYS 709  
 QY 232 ALALEEALKGPTDGGOSLEELSSHFTVIPHNFHSGQPPINSPPELQAKMDLVLAD 291  
 DB 710 ILSEVQAVSQGSDSQIL-DLSNRFYTLPHDFGMMKPPLLNADSVQAKVEMLDNLD 768  
 QY 292 IELAQAQAVSEQEKTYVEEYVPHPLDRDYQLLKCQLQLDLSGAPEYKVIQTYLEOT-GSNH 350  
 DB 769 IEVAYSLRGSDSSKSD----PIDVNYEKLKTDIKVVDKRDSEAEVIRKYVKNTHATT 824  
 QY 351 RCPQLQ--HWKVNQEGEEDRFQAHSKLGNRKLILWHTNMVAVNAILTSLRIMPH---- 404  
 DB 825 NAYDLEVMDFKIEREGESORYKPKQLHNRLHLWHSRTNFAGILUSQGLRIAPPAV 884  
 QY 405 SGRVKGKGIYFASSENSKAGYVIGMKCAHH---VGYMFLGEVALGREHINTDNPISLS 461  
 DB 885 TGYMFGKGIYFADVMVSKSANY-----CHTSQGDPIGLILLGEVALGNMYELKHAS-HISK 938  
 QY 462 PPPGFDSVIARGHTEPPTQDTLELDGQGVVPPQGPVPCPEFSSSTFSQSEYLIQOES 521  
 DB 939 LPRGKHSVKGLGKTPDPSPASITL-----GVEVPLGTGIP-SGVNDRCLLYNEYIVYDIA 993  
 QY 522 QCLRILYLLV 531  
 DB 994 QVNLKYLLKL 1003  
 RESULT 9  
 ID PPOL\_HUMAN STANDARD; PRT; 1013 AA.  
 AC P09874;  
 DT 01-MAR-1989 (Rel. 10, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
 DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1).  
 GN ADPRT OR PPOL OR PARP1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90091744; PubMed=2513174;  
 RA Auer B., Nagl U., Herzog H., Schneider R., Schweiger M.;  
 RT "Human nuclear NAD+ ADP-ribosyltransferase(polymerizing):  
 RT organization of the gene";  
 RL DNA 8:575-580(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=88076933; PubMed=3120710;  
 RA Uchida K., Morita T., Sato T., Ogura T., Yamashita R., Noguchi S.,  
 RT "Nucleotide sequence of a full-length cDNA for human fibroblast  
 RT poly(ADP-ribose) polymerase.";  
 RL Biochem. Biophys. Res. Commun. 148:617-622(1987).

RP SEQUENCE FROM N.A.  
RC TISSUE-Fibroblast; PubMed-2824474;  
RX MEDLINE-88058958; PubMed-2824474;  
RA Kurosaki T., Ushiro H., Mitsuuchi Y., Suzuki S., Matsuda M.,  
RA Matsuda Y., Katunuma N., Kangawa K., Matsuo H., Hirose T.,  
RA Inayama S., Shizuta Y.;  
RT "Primary structure of human poly(ADP-ribose) synthetase as deduced  
RT from cDNA sequence.";  
RL J. Biol. Chem. 262:15990-15997(1987).  
RN [14]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-88068596; PubMed-2891139;  
RA Cherney B.W., McBride O.W., Chen D., Alkhatib H., Bhatia K.,  
RA Hensley P., Smulson M.E.;  
RT "cDNA sequence, protein structure, and chromosomal location of the  
RT human gene for poly(ADP-ribose) polymerase.";  
RL Proc. Natl. Acad. Sci. U.S.A. 84:8370-8374(1987).  
RN [15]  
RP SEQUENCE OF 440-1013 FROM N.A.  
RX MEDLINE-87298455; PubMed-3113420;  
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,  
RA Miwa M.;  
RT "Molecular cloning of cDNA for human poly(ADP-ribose) polymerase and  
RT expression of its gene during HL-60 cell differentiation.";  
RL Biochem. Biophys. Res. Commun. 146:403-409(1987).  
RN [16]  
RP ERRATUM.  
RA Suzuki H., Uchida K., Shima H., Sato T., Okamoto T., Kimura T.,  
RA Miwa M.;  
RL Biochem. Biophys. Res. Commun. 148:1549-1550(1987).  
RN [17]  
RP SEQUENCE OF 1-94 FROM N.A.  
RX MEDLINE-91099327; PubMed-2125269;  
RA Yokoyama Y., Kawamoto T., Mitsuuchi Y., Kurosaki T., Toda K.,  
RA Ushiro H., Terashima M., Sumimoto H., Kuribayashi I., Yamamoto Y.,  
RA Maeda T., Ikeda H., Sagara Y., Shizuta Y.;  
RT "Human poly(ADP-ribose) polymerase gene. Cloning of the promoter  
RT region.";  
RL Eur. J. Biochem. 194:521-526(1990).  
RN [18]  
RP SEQUENCE OF 1-39 FROM N.A.  
RX MEDLINE-90211250; PubMed-2108670;  
RA Ogura T., Nyunoya H., Takahashi-Masutani M., Miwa M., Sugimura T.,  
RA Esumi H.;  
RT "Characterization of a putative promoter region of the human  
RT poly(ADP-ribose) polymerase gene: structural similarity to that of  
RT the DNA polymerase beta gene.";  
RL Biochem. Biophys. Res. Commun. 167:701-710(1990).  
RN [19]  
RP SEQUENCE OF 1-39 FROM N.A.  
RA Herzog H., Schneider R., Hirsch-Kauffmann M., Schnitzler D.,  
RA Schweiger M.;  
RL Submitted (JUL-1991) to the EMBL/GenBank/DBJ databases.  
RN [10]  
RP ANALYSIS OF ZINC FINGERS.  
RX MEDLINE-90222155; PubMed-2109322;  
RA Gradwohl G., Menissier de Murcia J., Molinete M., Simonin F.,  
RA Koken M.H.M., Hoeljmakers J.H.J., de Murcia G.M.;  
RT "The second zinc-finger domain of poly(ADP-ribose) polymerase  
RT determines specificity for single-stranded breaks in DNA.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:2990-2994(1990).  
RN [11]  
RP ANALYSIS OF ZINC FINGERS.  
RX MEDLINE-91072398; PubMed-2123876;  
RA Ikellima M., Noguchi S., Yamashita R., Ogura T., Sugimura T.,  
RA Gill D.M., Miwa M.;  
RT "The zinc fingers of human poly(ADP-ribose) polymerase are  
RT differentially required for the recognition of DNA breaks and nicks  
RT and the consequent enzyme activation. Other structures recognize  
RT intact DNA.";  
RL J. Biol. Chem. 265:21907-21913(1990).  
RN [12]  
RP ANALYSIS OF CATALYTIC DOMAIN.  
RX MEDLINE-91035460; PubMed-2121735;  
RA Simonin F., Menissier de Murcia J., Poch O., Muller S., Gradwohl G.,  
RA Molinete M., Penning C., Keith G., de Murcia G.M.;  
RT "Expression and site-directed mutagenesis of the catalytic domain of  
RT human poly(ADP-ribose) polymerase in *Escherichia coli*. Lysine 893 is  
RT critical for activity.";  
RL J. Biol. Chem. 265:19249-19256(1990).  
RN [13]  
RP NUCLEAR LOCALIZATION SIGNAL.  
RX MEDLINE-92371433; PubMed-1505517;  
RA Schreiber V., Molinete M., Boeuf H., de Murcia G.M.,  
RA Menissier de Murcia J.;  
RT "The human poly(ADP-ribose) polymerase nuclear localization signal is  
RT a bipartite element functionally separate from DNA binding and  
RT catalytic activity.";  
RL EMBO J. 11:3263-3269(1992).  
RN [14]  
RP MUTAGENESIS OF CATALYTIC DOMAIN.  
RX MEDLINE-97461532; PubMed-9315851;  
RA Rolli V., O'Farrell M., Menissier-de Murcia J., de Murcia G.M.;  
RT "Random mutagenesis of the poly(ADP-ribose) polymerase catalytic  
RT domain reveals amino acids involved in polymer branching.";  
RL Biochemistry 36:12147-12154(1997).  
CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYLATION). THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =  
CC nicotinamide + [ADP-D-ribosyl](N-1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBUNIT: HOMODIMER (Potential).  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; X16674; CAA34663.1; -;  
DR EMBL; M18112; AAA60137.1; -;  
DR EMBL; J03473; AAB59447.1; -;  
DR EMBL; M17081; AAA51599.1; ALT\_SEQ.  
DR EMBL; M32721; AAB60155.1; -;  
DR EMBL; M29786; AAA51663.1; -;  
DR EMBL; M29545; AAA51663.1; JOINED.  
DR EMBL; M29766; AAA51663.1; JOINED.  
DR EMBL; M29767; AAA51663.1; JOINED.  
DR EMBL; M29768; AAA51663.1; JOINED.  
DR EMBL; M29769; AAA51663.1; JOINED.  
DR EMBL; M29770; AAA51663.1; JOINED.  
DR EMBL; M29771; AAA51663.1; JOINED.  
DR EMBL; M29772; AAA51663.1; JOINED.  
DR EMBL; M29773; AAA51663.1; JOINED.  
DR EMBL; M29774; AAA51663.1; JOINED.  
DR EMBL; M29775; AAA51663.1; JOINED.  
DR EMBL; M29776; AAA51663.1; JOINED.  
DR EMBL; M29777; AAA51663.1; JOINED.  
DR EMBL; M29778; AAA51663.1; JOINED.  
DR EMBL; M29779; AAA51663.1; JOINED.  
DR EMBL; M29780; AAA51663.1; JOINED.





DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; Znf-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS0172; BRCT; 1.  
DR PROSITE; PS00347; PARP\_2N\_FINGER\_1; 2.  
DR PROSITE; PS00664; PARP\_2N\_FINGER\_2; 2.  
KW transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc.  
FT INT\_MET 0 0 BY SIMILARITY.  
FT DNA\_BIND 1 372 .  
FT DOMAIN 385 461 BRCT.  
FT DOMAIN 373 523 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 524 1013 NAD-BINDING.  
FT ZN\_FING 20 55 PARP-TYPE.  
FT ZN\_FING 124 161 PARP-TYPE.  
FT DOMAIN 206 208 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT MOD\_RES 1 1 BLOCKED (BY SIMILARITY).  
FT MOD\_RES 407 407 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 413 413 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 435 435 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 437 437 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 444 444 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 456 456 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 484 484 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 488 488 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 491 491 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 512 512 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 513 513 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 519 519 ADP-RIBOSYL[N] (POTENTIAL).  
FT CONFLICT 638 638 Y -> H (IN REF. 4).  
FT CONFLICT 641 641 E -> A (IN REF. 4).  
FT CONFLICT 752 752 N -> D (IN REF. 4).  
SQ SEQUENCE 1013 AA; 112529 MW; AA366F2B29BE97C0 CRC64;

Query Match 23.6%; Score 665.5; DB 1; Length 1013;  
Best Local Similarity 33.58; Pred. No. 9.1e-39;  
Matches 186; Conservative 99; Mismatches 203; Indels 67; Gaps 23;  
QY 1 MAPKPKPVQTEGPEKKKGROAGREDDPFRSTAEAKAIPAEKRI-----IRVDPTCP 53  
DB 493 VVPKGGK-----SAAPSKSKSGAVKEGVNKS-----EKRKKLTGKGAADVDPDSG 537  
QY 54 LSSNPGTOVED-----YNTLNTQNIENNKNFYIIQLQ--DSNFFTCWNWGRVGEV- 107  
DB 538 LEHS--AHVLEKGGKVFSAITGLVDIVGTNTSYKQLLESDDKESRYWFRSGRGVTI 595  
QY 108 GOSKINHTRLEDADKDFEKKFKETKKNWAERDFVSHPGKTYLTIEVQ--AEDEAQEAVV 166  
DB 596 GSNLEQMPQSKEDAVEHFMKLYTEETGTNAHWSK-NFTYIPKRFYPLEIDYGDG--EAVK 652  
QY 167 KVDGPGVTRTVKRVQPCSLDPATQKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKOQI 226  
DB 653 KL--AVKPGTK-----SKLPKPVQELVGMFEDVESMKKALVEYIDLQKMLPLGKLSRQI 705  
QY 227 ARGFEALAEALGKPTDGGOSLELSHFYTVIPIPHNGHSQPPPIPSPELLQAKKMDL 286  
DB 706 QAAQSILSEVOQAVSQGSSQIL-DLSNREYTLIPHDFGKKPPLLNNTDSVQAKVEML 764  
QY 287 LVLDIELAQAQAVSEQKVEEVPHPHLDYQLLKQLQQLDLSGAPPEYKVIQTYLQST 346  
DB 765 DNLDLIEVAYSLLRGGSDSSKD-----PIDVNYEKLKTDIKVDRDSEAEVIRKYVKT 820  
QY 347 -GSNHRCPQLQ--HIWKVNOGEEDRFQAHSKLGNRKLWLHGTNMVVAAILTSGLRIMP 403

DB 821 HATTHNAYDLEVIDIFKIEREGESQRYKPFQRLNRRLLWHGSRRTNFAGILSQGLRIAP 880  
QY 404 H-----SGRGVKGIGYFASSENSKAGYVIGMKCGAHH---VGYMELGEVALGREHINTDN 456  
DB 881 PEAPVTGYMFGKGIYFADVMVSKSANT-----CHTSOGDPGIGILLGEVALGNMYELKHAS 935  
QY 457 PSLKSPPPGFDVSVIARGHTEPDPTQDTELELDGQVVPQGVPCPEPSSSTFSQSEYL 516  
DB 936 -HISKLPKKGKSHVKGKGTAPDPS--ASITLDG--VEVPLGTGIP-SGVNDICLLYNEYI 989  
QY 517 IYQESQCLRYLLEY 531  
DB 990 VYDIAQVNLKYLKL 1004  
RESULT 11  
PPOL\_XENLA  
ID PPOL\_XENLA STANDARD; PRT; 998 AA.  
AC P31669;  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)-ADP-  
riboseyltransferase) (Poly[ADP-ribose] synthetase) (Fragment).  
OS Xenopus laevis (African clawed frog).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;  
OC Xenopodinae; Xenopus.  
OX NCBI\_TaxID=8355;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Ovary;  
RA Saulier-Le Drian B.M.;  
RL Thesis (1992), University of Rennes, France.  
RN [2]  
RP SEQUENCE OF 742-876 FROM N.A.  
RX MEDLINE-93277538; PubMed-8503897;  
RA Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,  
RA Miwa M.;  
RT "Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)  
polymerase from Xenopus laevis and cherry salmon using heterologous  
oligonucleotide consensus sequences.";  
RL Biochem. Biophys. Res. Commun. 193:119-125(1993).  
CC -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
PROTEINS BY POLY[ADP-RIBOSYLATION]. THE MODIFICATION IS DEPENDENT  
ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor -  
nicotinamide + (ADP-D-riboseyl)(N+1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,  
AND BRAIN. LOW IN LIVER.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
-----  
DR EMBL; 212139; CAA78126.1; -.  
DR EMBL; D13810; BRA02966.1; -.

DR PIR; S31735; S31735.  
 DR HSP; P26446; 1A26.  
 DR InterPro; IPR001357; BRCT.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR001510; Znf-PARP.  
 DR Pfam; PF00533; BRCT; 1.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF00645; zf-PARP; 2.  
 DR ProDom; PD004675; Znf-PARP; 2.  
 DR SMART; SM00292; BRCT; 1.  
 DR PROSITE; PS0172; BRCT; 1.  
 DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
 DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
 KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 KW ADP-ribosylation; Zinc-finger; Zinc.  
 FT NON\_TER 1 1  
 FT DNA\_BIND <1 356  
 FT DOMAIN 357 507 AUTOMODIFICATION DOMAIN.  
 FT DOMAIN 357 507 BRCT.  
 FT DOMAIN 369 445 NAD-BINDING.  
 FT DOMAIN 508 998 PARP-TYPE.  
 FT ZN\_FING 8 43 PARP-TYPE.  
 FT ZN\_FING 111 148 PARP-TYPE.  
 FT DOMAIN 193 195 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
 FT DOMAIN 207 212 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
 FT MOD\_RES 391 391 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 397 397 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 419 419 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 428 428 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 429 429 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 447 447 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 454 454 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 467 467 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 471 471 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 477 477 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 495 495 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 496 496 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 503 503 ADP-RIBOSYL[N] (POTENTIAL).  
 FT CONFLICT 746 746 Q -> E (IN REF. 2).  
 SQ SEQUENCE 998 AA; 111126 MW; F5A25E4A3366BAE7 CRC64;

Query Match 23.1%; Score 653; DB 1; Length 998;  
 Best Local Similarity 32.5%; Pred. No. 6.6e-38;  
 Matches 177; Conservative 102; Mismatches 214; Indels 52; Gaps 20;

QY 3 PKPKPWVOTEGEKKKGKQAGREEDPFRSTAEALKAIPAERIRVDPCTPLSNPQTQV 62  
 DB 480 PSSGPVAGKSGKKEKSGKSEKMKLVKGGAAIDPDEL---EDSCHVLETTG--- 533  
 QY 63 YEDYNTLNQTNENNKKFVLIQLQ-DSNRFCTCNWRGVRCEV-QQSKINIFTRLED 120  
 DB 534 -KIFSATLGLVDITRGNSYTKQLIEHDSRVYWRFSRGRVGTGVSCKLEEMSSKED 592  
 QY 121 AKKDFEKKFRKTKNNNAERDHFVSHPGKTYLIEQVQAEQAEVAVKVRGVPVTRKRV 180  
 DB 593 AIEHFLNLYQDQTNAM-HSPNFTKYPKKYPLEIDYQOE-EDVVKLSVG-AGTKSKLA 649  
 QY 181 QPCSLDPATQKLITNIFSKEMKNTMALMDLDVKKMPLGKLSKQOIQARGFALPALEAL 240  
 DB 650 KP-----VQELIKLIFDVESMKAMVFEIDLQKMPGLGSKRQIQSAVSILSQVQAV 703  
 QY 241 KGPTDGGQSLEE-----LSSHFFVIVPHNFGHSOPPPINSPELLOAKMDLLVLADI 295  
 DB 704 -----SESLSEARLLDLSNQFYTLIPHDFGKPKPPPLNNLEYIQAKVQMDNLLDIEVA 757  
 QY 296 QAL--QAVSEQEKTVVEVPHPLDRDYOLLKCOLLDSDGAPEYKVIOTYLEOT-GSNHRC 352  
 DB 758 YSLRGADGGEK-----DPIDVYKEIKITDIAKVAKDESRILIDYKNTHTADTHNA 811  
 QY 353 PTLQ--HIWKVNOGEEDRFQAHSKLGNRKLHLLHGHNTNMAVVAAILTSGLRIMPH---SG 406

DB 812 YDLEVLFIKIDREGEYQRYKPKQLHNRQLLWHGSRRTNFAGILSQGLRIAPPEAVTG 871  
 QY 407 GRVCKGIYFASSENSKSGAGYVIGMKCGAHVGYMELGEVALGREGHHINTDNPSLKSPPGF 466  
 DB 872 YMFCKGIYFADVMVSKSANYCHAMP--GSPGLILLGEVALGNMHELMKAASOITKL-PK 928  
 QY 467 DSVIARGHTPDPTQDTELELDGQVVVQGPVPCPEFSSSTFSQSEYLIYQESQCLRLR 526  
 DB 929 HSKVGLGRTAPDPS--ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIYVDIAQVNLK 983  
 QY 527 YLLEV 531  
 DB 984 YLLKL 988  
 RESULT 12  
 PPOL\_MOUSE  
 ID PPOL\_MOUSE STANDARD; PRT; 1012 AA.  
 AC P11103; Q9JLK4; Q9QVQ3;  
 DT 01-JUL-1989 (Rel. 11, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+)  
 DE ADP-ribosyltransferase-1) (Poly[ADP-ribose] synthetase-1) (msPARP).  
 GN ADPRT OR ADPRT1 OR ADPRP.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM 1).  
 RX STRAIN=BXSB;  
 RX MEDLINE=89263780; PubMed=2498841;  
 RA Huppi K., Bhatia K., Siwarski D., Klinman D., Cherney B., Smulson M.;  
 RT "Sequence and organization of the mouse poly (ADP-ribose) polymerase  
 RT gene.";  
 RL Nucleic Acids Res. 17:3387-3401(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.  
 RC STRAIN=129/Sy X C57BL/6; TISSUE=Fibroblast;  
 RX MEDLINE=20270268; PubMed=10809783;  
 RA Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;  
 RT "Characterization of sPARP-1. An alternative product of PARP-1 gene  
 RT with poly(ADP-ribose) polymerase activity independent of DNA strand  
 RT breaks.";  
 RL J. Biol. Chem. 275:15504-15511(2000).  
 RN [3]  
 RP KNOCK-OUT.  
 RX MEDLINE=96007847; PubMed=7578427;  
 RA Auer B., Flick K., Wang Z.-Q., Haidacher D., Jaeger S., Berghammer H.,  
 RA Kofler B., Schweiger M., Wagner E.F.;  
 RT "On the biological role of the nuclear polymerizing NAD+ protein (ADP-  
 RT ribosyl) transferase (ADPRT): ADPRT from Dictyostellium discoideum and  
 RT inactivation of the ADPRT gene in the mouse.";  
 RL Biochimie 77:444-449(1995).  
 CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =  
 CC nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.  
 CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -1- SUBUNIT: HOMODIMER (Potential).  
 CC -1- SUBCELLULAR LOCATION: Nuclear.  
 CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a  
 CC short form/sPARP-1; may be produced by alternative initiation.  
 CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN



RA Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;  
RT "Genomic organization of Drosophila poly(ADP-ribose) polymerase and  
RT distribution of its mRNA during development.";  
RL J. Biol. Chem. 273:11881-11886(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Ananidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton R.G., Wortman J.R., Randell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.H., Helt G., Nelson C.K., Miklos G.L.G.,  
RA Abriel J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Bakendell J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J.J., Brokstein P., Brotter P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith H.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY[ADP-RIBOSYLATION]. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -1- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor =  
CC nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.  
CC -1- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM: ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES  
CC OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM  
CC IN LATER EMBRYOS.  
CC -1- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN  
CC EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.  
CC -1- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -1- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC  
CC EMBL: D13806; BAA02964.1; -;  
DR EMBL: AF051548; AAC24518.1; -;  
DR EMBL: AF051544; AAC24518.1; JOINED.  
DR EMBL: AF051545; AAC24518.1; JOINED.  
DR EMBL: AF051546; AAC24518.1; JOINED.  
DR EMBL: AF051547; AAC24518.1; JOINED.  
DR EMBL: AF002935; AAF45400.1; -;  
DR EMBL: AF002666; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: AF002892; AAF45445.2; ALT\_SEQ.  
DR PIR: A47474; A47474.  
DR HSP: P26446; I426.  
DR FlyBase: Fbgn010247; Parp.  
DR InterPro: IPR001357; BRCT.  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR004102; PARP\_reg.  
DR InterPro: IPR001510; Znf-PARP.  
DR Pfam: PF005333; BRCT; 1.  
DR Pfam: PF006444; PARP; 1.  
DR Pfam: PF02877; PARP\_reg; 1.  
DR Pfam: PF00645; zfp-PARP; 2.  
DR ProDom: PD004675; Znf-PARP; 2.  
DR SMART: SM00292; BRCT; 1.  
DR PROSITE: PS00172; BRCT; 1.  
DR PROSITE: PS00347; PARP\_ZN\_FINGER\_1; 1.  
DR PROSITE: PS00664; PARP\_ZN\_FINGER\_2; 2.  
KW Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.  
FT DNA\_BIND 1 367  
FT DOMAIN 368 507 AUTOMODIFICATION DOMAIN.  
FT DOMAIN 380 454 BRCT.  
FT DOMAIN 508 994 NAD-BINDING.  
FT ZN\_FING 19 54 PARP-TYPE.  
FT ZN\_FING 123 161 PARP-TYPE.  
FT DOMAIN 203 210 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
FT DOMAIN 228 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
FT VARSPIC 376 564 MISSING (IN SHORT ISOFORM).  
SQ SEQUENCE 994 AA; 113791 MW; ACA85A270DD29E08 CRC64;  
  
Query Match 22.6%; Score 638; DB 1; Length 994;  
Best Local Similarity 32.2%; Pred. No. 7.4e-37;  
Matches 168; Conservative 100; Mismatches 195; Indels 58; Gaps 19;  
  
QY 37 KAIPAERKI-----IRVDPTCLSSNPGTQVED----YNCNTLNQNTIENNNKFFYIQL 87  
DB 500 KMPVSTFKVKDGLAVDPDSGLDI--AHVYVDSNNKYSVGLGLTIQRKNKSYKVL 557  
QY 88 LQ-DSNRFFTCWNRWVG-EVGOSKINHFTRLDEAKDKFEKFKREKTKNNNAERDFVS 145  
DB 558 LKADKKKEKYIFRSWGRIGTGNIGNSKLEEDTSESARKNEKEIYADKTGNEYEDNFVK 617  
QY 146 HPCKYTLIEVQAEDEAQAQVAVVVDGPRVTVTKRVQPCSLDPATOKLITIFSKEMFKNT 205  
DB 618 RTGRMYPTEIQYDD--QKLVKHE-----SHFFTSKLEISVQNLIKLIIFIDSMNKT 667  
QY 206 MALMDLVKMKGLSKQKQIARGFEALAEALGKPTDGGOSLEELSHFTVTIPHPNF 265  
DB 668 LMEFHIDMDKMPGLGKLSAQIQSVYRVKVIYNVLECGSNWAK-LIDATNRYFTLIPNF 726  
QY 266 GHSQPPPIINPELLQAKKMLLVADIETLAQALQAVSEKTEVEVPHPLDRDYQLLKQC 325  
DB 727 GVQLPTLIETHQOIEDLRQMLDSLAEITVAYSI----IKSEVDSDACNPLDNHYAQIKTQ 782  
QY 326 LQLDSDGAPYKVIQNYLEOT-CGNHRCPTLQ--HTWKVNOEGEEDRFQAHSLGNKLL 382  
DB 783 LVALDKNEEFSILSQYVKNHASTHKSVDLKIYDVFVKSQGEARRFKPFKLLHNRKLL 842





**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:01:28 ; Search time 117.38 seconds  
(without alignments)  
785.537 Million cell updates/sec

Title: US-09-701-586B-4  
Perfect score: 2823  
Sequence: 1 MAPKPKWVQTGEPKKGK.....EVLIIQESQRLRYLLEVLH 533

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_19.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archheap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2823	100.0	533	4	Q96CG2
2	2249.5	79.7	528	11	Q9LYR6
3	826	29.3	612	5	Q9TX06
4	694	24.6	635	10	O81294
5	670.5	23.8	653	10	O50017
6	669.5	23.7	607	13	Q9PS82
7	669.5	23.7	1014	11	Q921K2
8	636.5	22.1	607	13	Q9PS81
9	636.5	22.5	945	5	Q9N4H4
10	629	22.3	593	5	Q9TX05
11	598.5	21.2	983	10	Q9ZP54
12	598.5	21.2	1009	10	Q9SUT4
13	578	20.5	969	10	O24570
14	578	20.5	980	10	Q9ZSV1
15	534	18.9	727	5	Q9XUA5
16	306.5	10.9	2276	5	Q9TXQ1

17	302	10.7	815	10	Q9SWB4
18	284	10.1	815	10	Q9FK91
19	138.5	4.9	1327	4	Q9S271
20	137	4.9	1166	4	Q9H2K2
21	137	4.9	1265	4	Q9HAS4
22	134	4.7	1156	16	O66878
23	127	4.5	363	4	Q9H8R9
24	125	4.4	1181	5	Q9X237
25	125	4.4	1181	5	Q9VBP3
26	124.5	4.4	429	11	Q924M2
27	122.5	4.3	848	4	Q9Y2K0
28	120.5	4.3	419	4	Q9H9X9
29	120	4.3	1219	5	Q9NJ23
30	120	4.3	1229	5	Q9NJ22
31	120	4.3	1243	5	Q9NJ21
32	120	4.3	1253	5	Q9NJ20
33	120	4.3	1951	5	Q17042
34	119.5	4.2	927	11	Q91V71
35	119.5	4.2	946	11	Q91XT8
36	119.5	4.2	1021	11	Q91XU9
37	119.5	4.2	1052	11	Q91XT7
38	119	4.2	2167	10	Q9SS01
39	119	4.2	2182	10	Q9LW97
40	117	4.1	261	4	Q9H8F2
41	116.5	4.1	1435	3	Q03291
42	116	4.1	429	2	Q9LAX7
43	115.5	4.1	523	11	Q9DA63
44	115.5	4.1	744	4	Q9BXY9
45	114.5	4.1	1214	10	Q9FHI5

ALIGNMENTS

RESULT 1

Q96CG2 ID Q96CG2 PRELIMINARY; PRT; 533 AA.  
AC Q96CG2:  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 60.1 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_faxid=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=PRIMARY B-CELLS FROM TONSILS;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014260; AAI14260.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 533 AA; 60089 MW; 6296A0E439CC7767 CRC64;

Query Match	100.0%;	Score	2823;	DB	4;	Length	533;
Best Local Similarity	100.0%;	Pred.	No. 2.3e-208;				
Matches	533;	Conservative	0;	Mismatches	0;	Indels	0;
Gaps	0;						
Qy	1	MAPKPKWVQTGEPKKGKRGROAGREEDPFRSTAALKAIPAEKRIIRVDPTCPSSNPGT	60				
Db	1	MAPKPKWVQTGEPKKGKRGROAGREEDPFRSTAALKAIPAEKRIIRVDPTCPSSNPGT	60				
Qy	61	QVYEDYNCTLQNTNIENNKKFYIIQLLQDSNRFFTCNWRGRVGEVQSQKINHTRLD	120				
Db	61	QVYEDYNCTLQNTNIENNKKFYIIQLLQDSNRFFTCNWRGRVGEVQSQKINHTRLD	120				
Qy	121	AKKDFEKKFKREKTKNNNAERDHFVSHPGKYTLIEVQAEDEAQAQVVKVDRGPFVTVTKRV	180				
Db	121	AKKDFEKKFKREKTKNNNAERDHFVSHPGKYTLIEVQAEDEAQAQVVKVDRGPFVTVTKRV	180				
Qy	181	QPCSLDPATQKLTNITSKEMFKNTMALMDLVKKMPLGKLSKQOIARGFEALEEAL	240				

```
Db 181 QPCSLDPATQKLTNTIFSKEMFNKNTALMDLVKKPLGKLSKQIARGFEALEALEAL 240
QY 241 KGPTDGGQSLLEELSSHYFTVPIPNFGHSQPPPIINSELLOAKKMDLLVLADI LAQALQA 300
Db 241 KGPTDGGQSLLEELSSHYFTVPIPNFGHSQPPPIINSELLOAKKMDLLVLADI LAQALQA 300
QY 301 VSQEKTVVEVPHPLDRDYQLLKCQLQLDLSGAPEKVKIQTYLEOTGSHNRCPTLQHIWK 360
Db 301 VSQEKTVVEVPHPLDRDYQLLKCQLQLDLSGAPEKVKIQTYLEOTGSHNRCPTLQHIWK 360
QY 361 VNOGEDRQFAHSGKLGKLNRLWHGTMNAVVAAILTSGLRIMPHSGRGRVKGIFYASENS 420
Db 361 VNOGEDRQFAHSGKLGKLNRLWHGTMNAVVAAILTSGLRIMPHSGRGRVKGIFYASENS 420
QY 421 KSAGYVIGMKCGAHVGYMFLGEVALGREHHINTDPSLSPPPGDFSDSVIARGHTPDPT 480
Db 421 KSAGYVIGMKCGAHVGYMFLGEVALGREHHINTDPSLSPPPGDFSDSVIARGHTPDPT 480
QY 481 QDTELELDGQVVVPOGQVPVPCPEFSSSTFSQSEYLIYQESQCRRLYLLEVLH 533
Db 481 QDTELELDGQVVVPOGQVPVPCPEFSSSTFSQSEYLIYQESQCRRLYLLEVLH 533

RESULT 2
Q91YR6 PRELIMINARY; PRT; 528 AA.
ID Q91YR6 AC Q91YR6 PRELIMINARY; PRT; 528 AA.
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DE HYPOTHETICAL 59.4 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014870; AAH14870.1; -.
KW Hypothetical protein.
SQ SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;
```

```
Query Match 79.7%; Score 2249.5; DB 11; Length 528;
Best Local Similarity 80.7%; Pred. No. 2.4e-164;
Matches 431; Conservative 37; Mismatches 59; Indels - 7; Gaps 4;

QY 1 MAPKPKWQTEGPEKKKGROAGREEDPFRSTAEALKAIPAERKRIIRVDPCTPLSSNPGT 60
Db 1 MAPKRASVQTEG--SKKQKQTEEDSFRSTAEALRAAPADNRVIRVDPSCFSPRNPGI 58

QY 61 QVYEDYNTLNQNIENNKNFYIIQLQDSNRFFTCWNRWGRVGEVGSQKINHFTLED 120
Db 59 QVHEDYDCTLNQNIENNKNFYIIQLLEBSRFF--CWNRWGRVGEVGSQKMHFTLED 117

QY 121 AKDQFEKKFEKTKNNAERDHFVSHPGKTYLIEVQAEQAQVAVKVDGCPVTVTKRV 180
Db 118 AKDQFEKKFEKTKNNAERDHFVSHPGKTYLIEVQAEQAQVAVKVDGCPVTVTKRV 174

QY 181 QPCSLDPATQKLTNTIFSKEMFNKNTALMDLVKKPLGKLSKQIARGFEALEALEAL 240
Db 175 KPCSLDPATONLTNTIFSKEMFNKNTALMDLVKKPLGKLTQIARGFEALEALEAM 234

QY 241 KGPTDGGQSLLEELSSHYFTVPIPNFGHSQPPPIINSELLOAKKMDLLVLADI LAQALQA 300
Db 235 KNPSTDGQSLLEELSSHYFTVPIPNFGHSQPPPIINSELLOAKKMDLLVLADI LAQALQA 294

QY 301 V-SEQEKTVEVPHPLDRDYQLLKCQLQLDLSGAPEKVKIQTYLEOTGSHNRCPTLQHIW 359
Db 295 APGEEBEKVEVPHPLDRDYQLLRCQLQLDLSGESEYKAIQTLYLKQTGNSYRCPNLRHW 354
```

```
QY 360 KVNQEGEDRQFAHSGKLGKLNRLWHGTMNAVVAAILTSGLRIMPHSGRGRVKGIFYASEN 419
Db 355 KVNREGEDRQFAHSGKLGKLNRLWHGTMNAVVAAILTSGLRIMPHSGRGRVKGIFYASEN 414
QY 420 KSAGYVIGMKCGAHVGYMFLGEVALGREHHINTDPSLSPPPGDFSDSVIARGHTPDPT 479
Db 415 KSAGYVVTMHCGHGVGYMFLGEVALGREHHITIDPSLSPPPGDFSDSVIARGQTEPDP 474
QY 480 TDTELELDGQVVVPOGQVPVPCPEFSSSTFSQSEYLIYQESQCRRLYLLEVLH 533
Db 475 AQDIELELDGQVVVPOGQVPVPCPEFSSSTFSQSEYLIYQESQCRRLYLLEIHL 528

RESULT 3
Q9TX06 PRELIMINARY; PRT; 612 AA.
ID Q9TX06 AC Q9TX06 PRELIMINARY; PRT; 612 AA.
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
SQ SEQUENCE 612 AA; 69241 MW; CB340F7A88FF2364 CRC64;
```

```
Query Match 29.3%; Score 826; DB 5; Length 612;
Best Local Similarity 38.5%; Pred. No. 5.8e-55;
Matches 217; Conservative 79; Mismatches 194; Indels 74; Gaps 19;
```

```
QY 3 PRPKPWQTEGPEKKKGROAGREEDPFRSTAEALKAIPAERKRIIRV-----DPTCLP 54
Db 86 PTKKTKTNTSEAEK---TAASDLDDSSSESE-----DEKNOISVKIKGRAANDPHFPD 136

QY 55 SSNPGTQVYED---YNTLNQNIENNKNFYIIQLQ--DSNRFFTCWNRWGRVGEVQ 109
Db 137 SR---XHYVENGKDVYDNLNQTETSONNNKXYIIQLLEADGSSYVWNRWREGKLGQ 193

QY 110 SKINHTR--LEDAAKDFEKKFREKTKNNAERDHFVSHPGKTYLIEVQAEQAQVAVK 167
Db 194 SSRKDFGKGLNQAISLFCSEKFEKTKNTFDRANEKVKAGYKDYMIELDYSTD-----K 248

QY 168 VDRG----PVRTVTKRV---QPCSLDPATQKLTNTIFSKEMFNKNTALMDLVKKMPLG 219
Db 249 PKNGASTTATTTTKVVEHKEKESLDERVELVKLIDFKMMERTMTAEKYLKKMPLG 308

QY 220 KLSKQIARGFEALEALEALKGPTDGGQSLLEELSSHYFTVPIPNFGHSQPPPIINSP 279
Db 309 KLSKNOITKGYLVKQIEDVMGKS--GESLSTLSRFFYTIIPAFGMSVPPVINTNQML 366

QY 280 QAKKMDLLVLADI LAQALQAQVSEKQTEVEVPHPLDRDYQLLKCQLQLDLSGAPEYKI 339
Db 367 IEKMNMLQNLADIETATNIKDSDES-----NILEHYAKLKTIDIQPLDENSECEYKNI 421

QY 340 QYILEOTGSHNRCPTLQHIWKVNOEGEDRQFAHSGKLGKLNRLWHGTMNAVVAAILTSG 399
Db 422 LLYVKNYGGKKPTVNTVTFKIDRDGEADRYKTKKHLGNRKLMLWHGSRNTNYASIIISQGL 481
```

QY 400 RIMPH-----SGRGVKGIFYFASSENSKSYGVIGMKGAHHVGY-----MFLGEVALGREH 450  
 Db 482 RTAPEAPVSGYFVGKGYFADCMLSANSY-----CRTVGLDFCMGLLGDVALGKTA 534  
 QY 451 HINTDNPSLKSPPGFDSVIARGHTEPDT---QTELELDGQQVVPQGVPCPEFSS 507  
 Db 535 DLARDT-YNEKPQPNHSTWALGTVEPDKPVFFQDTE-----GCITPYGQMIPSOHKG 587  
 QY 508 STFSOSEYLIQESOCRLRYLLE 531  
 Db 588 SCY-EHQYVYVDVAOVHLKYLQL 610

RESULT 4  
 ID O81294 PRELIMINARY; PRT; 635 AA.  
 AC O81294;  
 DT 01-NOV-1998 (Tremblrel. 08, Created)  
 DT 01-NOV-1998 (Tremblrel. 08, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE T14P8.19 PROTEIN (NAD+ ADP-RIBOSYLTRANSFERASE).  
 GN T14P8.19 OR AT4G02390.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA WASHU;  
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Kalicki J., Elliott G., Cloud J.;  
 RT "The sequence of A. thaliana T14P8.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. COLUMBIA;  
 RA Waterston R.;  
 RL Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Lanar B., Stoneking T., Stumpf J., Mewes H.W., Lemcke K.,  
 RA Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF069298; AAC19283.1; -;  
 DR EMBL; AL161494; CAB80732.1; -;  
 DR HSSP; P26446; 1A26.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR003034; SAP.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF02037; SAP; 2.  
 DR SMART; SM00513; SAP; 2.  
 KW Transferase.  
 SQ SEQUENCE 635 AA; 72017 MW; E3F1CBE4D367A377 CRC64;

Query Match 24.6%; Score 694; DB 10; Length 635;  
 Best Local Similarity 34.2%; Pred. No. 8.3e-45;  
 Matches 185; Conservative 221; Indels 38; Gaps 18;  
 QY 11 TEGPEKKKGROAGREDFRSTAEALKAIPAEKRIIRV-DPTCPILSSNPGTQVYED---- 65  
 : | | | | : | | | | : | | | | : | | | | : | | | |

Db 107 SNAPVSSNDEADDNNGFEEKKEEKTIVTATKGAADVLDQWIPDEIKSOYHVLQRGDV 166  
 QY 66 YNCTLNQNIENNKKFYIIQLLQ-DSNRFTCNWRGVRGVEGOSKIN-HFTRLEDAKK 123  
 Db 167 YDAILNQTNRDNNKFFVLOVLESDSKTKYMYVTRWRGVGKSGKLDGPDYDSWDRAIE 226  
 QY 124 DFEKFRKTKNNWAERDHFVSHPGKTYLIEVOADEAQEAQVVKVDRGVRTVTKRVQP- 182  
 Db 227 IFTNKFNDKTNYSNDRKEFIPHPKSYTWLEMDYKEENDSPVND---IPSSSEYKPE 283  
 QY 183 -CSLDPATOKLITNIFSKEMFNKMTMALMDLVKMKPLGKLSKQOIARGFEALEEALK 241  
 Db 284 QSKLDRVAKFTSLICNVSMMAQHMIEGYNANKLPLGKISKSTISKGYEVLKRISVI- 342  
 QY 242 GPTDGGQSLLELSHFYTVIPHNFCHSQPPP--INSPELLOAKKMDMLLVLDIELAQALQ 299  
 Db 343 -DRYDRTRLEELSGEFYVIPHDFGFKMSQVITPQKLKOKIEMVEALGEIELATKLL 401  
 QY 300 AVSEQETVEEVPHPDRDYQLLKCOLQLLDGSAPEYKVIQTYLEQT-GSNHRCPTLQ-- 356  
 Db 402 SVDPGLQ-----DPLYHYQQLNCGLTPVGNDSSEFSWVANYMENTHAKTHSGYVTEA 456  
 QY 357 HIWKVQGEEDRFQAHKLNKRLNLTGHTNNAVVAAILTSLGRIMPH-----SGRGVKG 412  
 Db 457 QLFRAVRAVEADRFQOFSSKNRMLLHGSRLTNWAGILSQGLRIAPPEAPVTGYMFGKG 516  
 QY 413 IYFASNSKSGAGYVIGMKGAHHVGMFLGEVALGREHHINTDNPSLKSPPGFDSVIR 472  
 Db 517 VYFADFMSKSNYCYA-NTGAND-GVLLCEVALGDMNELLSDYNADNLPPGKLTSGV 574  
 QY 473 GHTEPDPTQDTELELDGQQVVPQGVV--PCPEFSSSTFSQSEYLIQESOCRLRYLLE 530  
 Db 575 GKTAPNPSEAQTLE-DG--VVVPLGKPVRSCT--SKGMLLYNEIYVYVNEQIKRYYIQ 628  
 QY 531 V 531  
 Db 629 V 629

RESULT 5  
 ID O50017 PRELIMINARY; PRT; 653 AA.  
 AC O50017;  
 DT 01-JUN-1998 (Tremblrel. 06, Created)  
 DT 01-JUN-1998 (Tremblrel. 06, Last sequence update)  
 DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
 DE POLY(ADP-RIBOSE) POLYMERASE.  
 GN PARP.  
 OS Zea mays (Maize).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
 OC Panicoideae; Andropogoneae; Zea.  
 OX NCBI\_TaxID=4577;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,  
 RA O'Farrell M., Van Montagu M., Inze D., Kushnir S.;  
 RT "Higher plants possess two poly(ADP-ribose) polymerases.";  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ222588; CAA10888.1; -;  
 DR HSSP; P26446; 1A26.  
 DR InterPro; IPR001290; PARP.  
 DR InterPro; IPR004102; PARP\_reg.  
 DR InterPro; IPR003034; SAP.  
 DR Pfam; PF00644; PARP; 1.  
 DR Pfam; PF02877; PARP\_reg; 1.  
 DR Pfam; PF02037; SAP; 2.  
 DR SMART; SM00513; SAP; 2.  
 SQ SEQUENCE 653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

Query Match 23.8%; Score 670.5; DB 10; Length 653;  
 Best Local Similarity 35.4%; Pred. No. 5.5e-43;

```
Matches 170; Conservative 93; Mismatches 186; Indels 31; Gaps 14;
QY 64 EDYCNLTNNTNNNNKFFYIIQLQ--DSNRFFTCNRRGRVGEVQSGKINHFTRLED-A 121
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 188 EYDAILNTQNVGNKKFYIIQVLESAGGSPWVYNNRGRVCRQDKLHGPSPTDQA 247
| : | | | | : | | | | | : | : | | | | | | | | : | : |
QY 122 KDFEKKFREKTNNNAERDHFVSHPGKYTLIEVQAEDEAAVAVKVDGRPVRTVTKRQ 181
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 248 IYFECKFHNTNNHNSDRKNPKYAKKYTWLEMDYGETEKE-----IEKG---SITDQIK 300
| : | | | | : | | | | | : | : | | | | | | | | : | : |
QY 182 PCSLDPATOKLTNIFSKEMFKNTMALMDLVKKMPLGKLSKQOQIARGFEALEALEAL 241
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 301 ETKLETRIAQFISLICINISMMKQWVEIGYNAEKLPLGLKATILKGYHLKRIISDVIS 360
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 242 GPTDGGSLSELSHFYTVIPHNFGHSOPPP--INSPELLOAKKMLLVLLADIELAQALQ 239
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 361 --KADRRHLEQLTGEYTVIPHDFGFRKREFIIDIPQKAKLEWVEALGEIEIATKL- 417
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 300 AVSEQKTVVEVPHPLDRDYQLLKCOLQLDLSGAPEYKVIQTYLEQT--GSNHRCP 356
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 418 ----LEDDSSDDQDPLVARYKQLHCDFTPLEADSDEYSMTKSYLNTHGKTHSGYTVDIV 473
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 357 HIWVQGEEDRFQASHKLGKRLKLLHGTNNMAVVAAILTSGLRIMPH-----SGGRVGK 412
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 474 QIEKVRHGETERFQFASRNRMLLHWSRLSNWAGILSQGLRIAPPEAPVTGYMFGKG 533
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 413 IYFASENKSAGYVIGMKGCAHVGVMFLGEVALGREHHINTONPSLKSPPGFDSVIAR 472
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 534 VYFADMFKSANTCYASE--ACRSVGLLICEVALGDMNELLNADYDANNLPGKLSKGV 591
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 473 GHTEPPTQDTLELDCQVVVPGQVPVCPPEFSSTFSQSEYLIYQESQRLRYLLEVH 532
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 592 GQTAPNWN-ESKVADG--VVVPLGEPKQEPS-KRGLLYNEVIVYNDQIRMRYVLHVN 647
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
RESULT 6
Q9PS82 PRELIMINARY; PRT; 607 AA.
AC Q9PS82;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauaria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD+: protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSP; P26446; I26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR PROSITE: PS0172; BRCT; 1.
SQ SEQUENCE 607 AA; 68033 MW; 75F6EED3D08F402 CRC64;
Query Match 23.7%; Score 669.5; DB 13; Length 607;
Best Local Similarity 33.2%; Pred. No. 5.9e-43;
Matches 186; Conservative 100; Mismatches 205; Indels 69; Gaps 24;
QY 7 PW---VOTEGPE-----KKKGQAGREDDPFRSTAELKAIAPAEKRIIRVDPT 51
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
```

```
Db 73 PNGAEVKEHQEVAVDVGKCSKPANMKSAGVKVEEQGPSKSEKKMKLV---KGAADVDP 129
| : | | | | | | | | | | : | : | : | : | : | : | : | : | : |
QY 52 CPLSSNPGTQVYED---YNCITLNTNIENNKKFYIIQLQD--SNRFFTCNRRGRV 105
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 130 SGLEDS--AHVFEKGGKIFSAITGLVDIVKGTNSYKLOLLEDDRESRYWV-FRSN 186
| : | | | | : | | | | | : | : | | | | | | | | : | : |
QY 106 EV-GQSKINHTRLEDAKKDFEKKFREKTNNNAERDHFVSHPGKYTLIEVQ-AEDE 163
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 187 TVIGSNKLEQMPSKEDAVEHFLNLYEKTGNSWHS-NEFKYKPKFYPLEIDYGQDEE 245
| : | | | | : | | | | | : | : | | | | | | | | : | : |
QY 164 AVKVDGRPVRTVTKRQVPCSLDPATOKLTNIFSKEMFKNTMALMDLVKKMPLGKLSK 223
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 246 RKLTVSAG--TKSLAKP-----IODLKMTIFVESMKAMVEFEIDLQKMPGLKLSK 296
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 224 QQTARCFEALAEALKAGTGGQSL-ELSSHFTYVIPHNFHSGHSQPPPIINSP 281
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 297 RQIQSAYSTILNEVOQAV---SDGSESQILDLSNRRFTYLPHPFGMKKPLLSNLE 353
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 282 KKDMLLVLLADIQAALQAVSEQKTVVEVPHPLDRDYQLLKCOLQLDLSGAPEYKVIQ 341
| : | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 354 KVQMLDNLDDIEVAYSILLRGNGEDGKD-----PIDINYEKLRDIDKVVDDKSEAK 409
| : | | | | : | : | : | : | : | : | : | : | : | : | : | : |
QY 342 YLEQT--GSNHRCP 398
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 410 YVKNTHAATHNAYDLKAVVEIFRIEREGESQRYKPFKOLHNRQLLHWSRTNFA 469
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 399 LRIMPH---SGGRVGKGIYFASENSKSAGYVIGMKG---GAHHVGVMFLGEVALGR 451
| : | | | | : | : | : | : | : | : | : | : | : | : | : | : |
Db 470 LRITAPPEAPVTGYMFGKGIYFRDMVSKSANY-----CHTSQADPILGILLGEVALGN 524
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 452 INTDNP 511
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 525 LKNASHITKL-PGKSHSVKGLGKTAPDPATT--TLDG--VEVPLNGI-STGIND 578
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
QY 512 QSEYLIYQESQRLRYLLEV 531
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 579 YNEYIVYDVAQVNLKYLLKL 598
| : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
RESULT 7
Q92IK2 PRELIMINARY; PRT; 1014 AA.
AC Q92IK2;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE SIMILAR TO ADP-RIBOSYLTRANSFERASE (NAD+, POLY (ADP-RIBOSE)
DE POLYMERASE).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC012041; AAH12041.1; -.
KW Transferase.
SQ SEQUENCE 1014 AA; 112721 MW; 5DCE68E4CB3F46EB CRC64;
Query Match 23.7%; Score 669.5; DB 11; Length 1014;
Best Local Similarity 33.5%; Pred. No. 1.3e-42;
Matches 186; Conservative 99; Mismatches 203; Indels 67; Gaps 22;
QY 1 MAPKPKPWOTEGPEKKKGQAGREDDPFRSTAELKAIAPAEKRI-----INVD 53
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 494 VAPKPK-----SAAPSKSKGAVKEGVNKS-----EKRMKLTGKGAADVDP 538
| : | | | | : | | | | | : | : | | | | | | | | : | : |
QY 54 LSSNPGTQVYED---YNCITLNTNIENNKKFYIIQLQD--SNRFFTCNRRGRV 107
| : | | | | : | | | | | : | : | | | | | | | | : | : |
Db 539 LEHS--AHVLEKGGKVFSAITGLVDIVKGTNSYKLOLLEDDRESRYWFSN 596
| : | | | | : | | | | | : | : | | | | | | | | : | : |
```





```

DR HSP; P26446; 1A26.
KW Transferase; Glycosyltransferase; NAD.
FT CHAIN 2 983 POLY(ADP-RIBOSE) POLYMERASE.
SQ SEQUENCE 983 AA; 111232 MW; 468E12A8EF1B6F4F CRC64;

Query Match 21.2%; Score 598.5; DB 10; Length 983;
Best Local Similarity 31.1%; Pred. No. 3.4e-37;
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;

QY 17 KKGROAGREDFRSTAEALKAIPAEKRIIRVDPTCLSSNPGTQ-----VYED----YNC 68
DB 472 KKORKLPDFKDIETSESLVTVKVRG-----SAVHEASGLQHCHILEDGNISYNT 524

QY 69 TLNQNIENNKKFYIQLQDSNRFCTW--NRWGRVG--EVGOSKINHTRLEDAKD 124
DB 525 TILMSDLSGTINSYIILQIQE--DKGSDCYVFRKGRVGNKIGGNKVEEMSK--SDAVHE 582

QY 125 FEKKFREKTKN---NWAERDHFVSPGKYTLIEVOAEDAQEAUVKVDGPGVTVTKRVQ 181
DB 583 FKRLFEKTKNTWESWEQKTFNQKQKFLPLD-----IDYGVNKQVAKK-E 628

QY 182 P---CSLDPAQKLTITNLFSEKEMKNTMALDVKMKPLGKLSKQOITARGFEALEALE 237
DB 629 PFQTSNLSAPSLIELMKMLFDVETYSAMFEINNSEMPGLKLSKHNTOKGFEALTEIQ 688

QY 238 EAL-----KGPTDGGOSLELSHFYTVIPHNFGHSOPPPINSPELIQAKKMDLLVLADIE 293
DB 689 RLTTESDPQTKESLLVDASNRFTMIP-----SIHPHIIREDDFKSKVKMLEALQDIE 744

QY 294 LAQALQAVSEQEKTVEEVPHPDRDYQLKQLQLDGAPEYKVIQTYLEQTSNHRCP 353
DB 745 IASRI--VGFVDVSTES----LDDKYKLCIDISPLPHDSEDYRLIEKYLNIT---HAP 794

QY 354 T-----LQHIWKVNOEGEEDRFQAH--SKLGNRKLLHGTNMAVVAAILTSGLRI---- 401
DB 795 TTETWSLEEEVFALEGEFQKAPHREKLGKMLLHGHGSRITNFGVILNQLGRIAPPE 854

QY 402 MPHSGRGVKGIFYFASENSKSAGYVIGMKCGAHVGYMFLGEALGREGHHINTDNP SLKS 461
DB 855 APATGYMFGKGIYFADLVSKSAQYC--YTCCKNPVGLMILLSEVALGEIHEL--TKAKYMDK 911

QY 462 PPPGSDSVIARGHTPEPDPTQDTELDGQGVVPGQVPVCPPEFSSTSSQSEYLIYQBS 521
DB 912 PPRGKHSTKGLGKKVP---QDSEFAKWRGDTVTPCGKPV--SSKVKASELMYNEYIYDFA 967

QY 522 QCRRLYLLEV 531
DB 968 QVKLOFLLV 977

RESULT 12
Q9SJW4 PRELIMINARY; PRT; 1009 AA.
AC Q9SJW4;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE PUTATIVE POLY (ADP-RIBOSE) POLYMERASE.
GN AT2G31320.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,

```

```

RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosomes 2 of the plant Arabidopsis
RT thaliana.";
RL Nature 402:761-768(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. COLUMBIA;
RX Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC006593; AAD20677.1; -.
DR HSP; P26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP-reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP-reg; 1.
DR Pfam; PF00645; Zf-PARP; 2.
DR ProDom; PD004675; Znf-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS50172; BRCT; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 2.
SQ SEQUENCE 1009 AA; 114133 MW; CDE6E41CC2A3A2DB CRC64;

Query Match 21.2%; Score 598.5; DB 10; Length 1009;
Best Local Similarity 31.1%; Pred. No. 3.5e-37;
Matches 171; Conservative 101; Mismatches 199; Indels 79; Gaps 23;

QY 17 KKGROAGREDFRSTAEALKAIPAEKRIIRVDPTCLSSNPGTQ-----VYED----YNC 68
DB 498 KKORKLPDFKDIETSESLVTVKVRG-----SAVHEASGLQHCHILEDGNISYNT 550

QY 69 TLNQNIENNKKFYIQLQDSNRFCTW--NRWGRVG--EVGOSKINHTRLEDAKD 124
DB 551 TILMSDLSGTINSYIILQIQE--DKGSDCYVFRKGRVGNKIGGNKVEEMSK--SDAVHE 608

QY 125 FEKKFREKTKN---NWAERDHFVSPGKYTLIEVOAEDAQEAUVKVDGPGVTVTKRVQ 181
DB 609 FKRLFEKTKNTWESWEQKTFNQKQKFLPLD-----IDYGVNKQVAKK-E 654

QY 182 P---CSLDPAQKLTITNLFSEKEMKNTMALDVKMKPLGKLSKQOITARGFEALEALE 237
DB 655 PFQTSNLSAPSLIELMKMLFDVETYSAMFEINNSEMPGLKLSKHNTOKGFEALTEIQ 714

QY 238 EAL-----KGPTDGGOSLELSHFYTVIPHNFGHSOPPPINSPELIQAKKMDLLVLADIE 293
DB 715 RLTTESDPQTKESLLVDASNRFTMIP-----SIHPHIIREDDFKSKVKMLEALQDIE 770

QY 294 LAQALQAVSEQEKTVEEVPHPDRDYQLKQLQLDGAPEYKVIQTYLEQTSNHRCP 353
DB 771 IASRI--VGFVDVSTES----LDDKYKLCIDISPLPHDSEDYRLIEKYLNIT---HAP 820

QY 354 T-----LQHIWKVNOEGEEDRFQAH--SKLGNRKLLHGTNMAVVAAILTSGLRI---- 401
DB 821 TTETWSLEEEVFALEGEFQKAPHREKLGKMLLHGHGSRITNFGVILNQLGRIAPPE 880

QY 402 MPHSGRGVKGIFYFASENSKSAGYVIGMKCGAHVGYMFLGEALGREGHHINTDNP SLKS 461
DB 881 APATGYMFGKGIYFADLVSKSAQYC--YTCCKNPVGLMILLSEVALGEIHEL--TKAKYMDK 937

QY 462 PPPGSDSVIARGHTPEPDPTQDTELDGQGVVPGQVPVCPPEFSSTSSQSEYLIYQBS 521
DB 938 PPRGKHSTKGLGKKVP---QDSEFAKWRGDTVTPCGKPV--SSKVKASELMYNEYIYDFA 993

QY 522 QCRRLYLLEV 531
DB 994 QVKLOFLLV 1003

```

## RESULT 13

024570 ID O24570 PRELIMINARY; PRT; 969 AA.  
AC O24570;  
DT 01-JAN-1998 (Tremblrel. 05, Created)  
DT 01-JAN-1998 (Tremblrel. 05, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE POLY(ADP-RIBOSE) POLYMERASE.  
GN PARP.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Babylchuk E., Cottrill P., Storozhenko S., Fungthong M.,  
RA O'Farrell M., Van Montagu M., Inze D., Kushnir S.;  
RT "Higher plants possess two poly(ADP-ribose) polymerases.";  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AJ222589; CAAL0889.1; -  
DR HSP; P26446; I426.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR003034; SAP.  
DR InterPro; IPR001510; ZnF-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; ZnF-PARP; 2.  
DR ProDom; PD004675; ZnF-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS0172; BRCT; 1.  
DR PROSITE; PS0064; PARP\_ZN\_FINGER\_2; 2.  
SQ SEQUENCE 969 AA; 109128 MW; EB23AC62EEC14009 CRC64;

Query Match 20.5%; Score 578; DB 10; Length 969;  
Best Local Similarity 30.6%; Pred. No. 1.2e-35;  
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

QY 16 KKKGRQAGREDDPFRSTAEALKAIPAERIRVDPCTPLSSNPGTVQYVEDYNCTLNQNI 75  
DB 480 KVKGRSAVHSSGLQDTAHILE-----DGKSI---YNATLMSDL 516  
QY 76 ENNNKFIYIOLL-QDSNRFTCNWRGVRG--EVGQSKINHFTLEDAKKDFEKKFK 132  
DB 517 ALGVSYYVLQIIQDDGSECYVFRKRWGVRGSEKIGGQKLEMSKTE-AIKEFKRLFLEK 575  
QY 133 TKNNNAE---RDHFVSHPGKTYTLIEVQAEDEAQAQVAVKVDGVPVTVTKRVQPCSLDPAT 189  
DB 576 TGNSEWECKTNFRKQGRFYPLDVG-----YGVKAPKPKDISEMK--SSLAPQL 625  
QY 190 OKLITNIFSKEMFNKNTMALMDLVKKMPLGKLSKQOQIARGFEALAEALKGPTDGGQS 249  
DB 626 LELMKLFNVETIYRAAMFEINSEMPLGKLSKENIEKGFALTEIQNLKLDKTDADQALA 685  
QY 250 LEE-----LSSHFTVIVPHNFGHSGOPPPINSPELLQAKKMLLVADIETLAQAQAV-SE 303  
DB 686 VRESLIVAASNRFFTLIP-----SIHPHIIRDEDDLMIAKAKMLEALQDIEIAKIVGFDSD 741  
QY 304 OETVEEVPHPLDRLDYOLKQLOLLDSDGAPEYKVIQTYLEQTSNHRCP-----LQ 356  
DB 742 SDES-----LDDKYMKLHCDITPLAHDSEYKLIQYL-----LNTHAPTHKWSLELE 790  
QY 357 HIWKYQREDEEDFOAH--SKLGNRKLHLWGTNMAVAAITLSGLRIMPH-----SGGRYVK 411  
DB 791 EVESLDRDGLNKSRYKNNLHNKMLLWHGSRSLTNFVGIISQGLRIAPPEAPVTGYMEGK 850  
QY 412 GIYFASNSKSGAGYVIGMKCA-----HHVGYMFLGEVALGREHHINTDNPISKSPPPGDS 468  
DB 851 GLYFADLYSKSAQY-----CYVDRNPNVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904

QY 469 VIARGHTPEPTQDTLELDGQVVVPOGVPVPCPEPFSSTFSOSEYLIYQESOCRLRYL 528  
DB 905 TKGLGKTVP---LESEFVKWRDDVVVPCGKVPV-SSIRSELMYNEYIYVNTSQVKMQFL 960  
QY 529 LEV 531  
DB 961 LKV 963

## RESULT 14

Q9ZSV1 ID Q9ZSV1 PRELIMINARY; PRT; 980 AA.  
AC Q9ZSV1;  
DT 01-MAY-1999 (Tremblrel. 10, Created)  
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).  
GN PARP1.  
OS Zea mays (Maize).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;  
OC Panicoideae; Andropogoneae; Zea.  
OX NCBI\_TaxID=4577;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE-99026291; PubMed-9808734;  
RA Mahajan P.B., Zuo Z.;  
RT "Purification and cDNA cloning of maize Poly(ADP-ribose) polymerase.";  
RL Plant Physiol. 118:895-905(1998).  
DR EMBL; AF093627; AAC79704.1; -  
DR HSP; P26446; I426.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR003034; SAP.  
DR InterPro; IPR001510; ZnF-PARP.  
DR Pfam; PF00533; BRCT; 1.  
DR Pfam; PF00644; PARP; 1.  
DR Pfam; PF02877; PARP\_reg; 1.  
DR Pfam; PF00645; ZnF-PARP; 2.  
DR ProDom; PD004675; ZnF-PARP; 2.  
DR SMART; SM00292; BRCT; 1.  
DR PROSITE; PS0172; BRCT; 1.  
DR PROSITE; PS0064; PARP\_ZN\_FINGER\_2; 2.  
KW Transferase; Glycosyltransferase.  
SQ SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;

Query Match 20.5%; Score 578; DB 10; Length 980;  
Best Local Similarity 30.6%; Pred. No. 1.3e-35;  
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

QY 16 KKKGRQAGREDDPFRSTAEALKAIPAERIRVDPCTPLSSNPGTVQYVEDYNCTLNQNI 75  
DB 491 KVKGRSAVHSSGLQDTAHILE-----DGKSI---YNATLMSDL 527  
QY 76 ENNNKFIYIOLL-QDSNRFTCNWRGVRG--EVGQSKINHFTLEDAKKDFEKKFK 132  
DB 528 ALGVSYYVLQIIQDDGSECYVFRKRWGVRGSEKIGGQKLEMSKTE-AIKEFKRLFLEK 586  
QY 133 TKNNNAE---RDHFVSHPGKTYTLIEVQAEDEAQAQVAVKVDGVPVTVTKRVQPCSLDPAT 189  
DB 587 TGNSEWECKTNFRKQGRFYPLDVG-----YGVKAPKPKDISEMK--SSLAPQL 636  
QY 190 OKLITNIFSKEMFNKNTMALMDLVKKMPLGKLSKQOQIARGFEALAEALKGPTDGGQS 249  
DB 637 LELMKLFNVETIYRAAMFEINSEMPLGKLSKENIEKGFALTEIQNLKLDKTDADQALA 696  
QY 250 LEE-----LSSHFTVIVPHNFGHSGOPPPINSPELLQAKKMLLVADIETLAQAQAV-SE 303  
DB 697 VRESLIVAASNRFFTLIP-----SIHPHIIRDEDDLMIAKAKMLEALQDIEIAKIVGFDSD 752



```

Qy 304 QKTVVEVPHPLDRDYQLKQQLQDLSGAPYKVIQTYLEQTGNSNRHCP------LQ 356
Db 753 SDES-----LDDYMKLHCDITPLAHSSEYKLIQYL-----LNTHAPTHKWSLE 801
Qy 357 HIWKVQGEEDRQAH-SKLGNRKLLWHGTNMAVAAILTSGLRIMPH-----SGGRVKG 411
Db 802 EVFSLDRGELNKYSRYKNNLHKLWHGSRLLNFVGLSQGLRIAPPEAPVTGYMFGK 861
Qy 412 GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGVALGREHINTDNPISLSPPPGFD 468
Db 862 GLYFADLYSKAAY-----CIYDRNPNVGLMLLSEVALGDMYELKAT-SMDKPPRGHS 915
Qy 469 VIARHTEPDPTQDTELELDQVQVVPQCPPEFSSSTFSOSELYIYQESQCLRYL 528
Db 916 TKGLGKTVP---LESEFVKWRDDVVVPCGKVP-SSIRSELMYNEYIVYNTSQVKMQFL 971
Qy 529 LEV 531
Db 972 LKV 974

RESULT 15
Q9XUA5
ID Q9XUA5 PRELIMINARY; PRT; 727 AA.
AC Q9XUA5;
DT 01-NOV-1999 (Tremblrel. 12, Created)
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE AC8.1 PROTEIN.
GN AC8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_taxid-6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83097; CAB05448.1; -.
DR HSSP; P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; znf-PARP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; znf-PARP; 1.
DR PRODOM; PD004675; znf-PARP; 1.
DR PROSITE; PS00064; PARP_ZN_FINGER_2; 1.
SQ SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFED CRC64;

Query Match 18.9%; Score 534; DB 5; Length 727;
Best Local Similarity 28.6%; Pred. No. 1.9e-32;
Matches 162; Conservative 104; Mismatches 211; Indels 90; Gaps 20;

Qy 30 RSTAAKAIAPAEKRIIRVDPCTPLSSNPGTQVYEDYNTLNQNIENNKKFYIQLLQ 89
Db 182 KATGEYVALAAGG--STEPATPASAP-TPPEAETVLSAEGSPSSNKRKPASAEIIE 238
Qy 90 -----DSNRFF---TCNWRGVRGEVQSKI-----NHFTLED 120
Db 239 IDGEGNPDENFAKRRMKKEARLMEVQKRMKKQSDLLWEYQIFERMPYTDNISILRE 298
Qy 121 AKKDF-----EKKFREKTKNNWAERDHFVSHPGKYTLIEVQAEADQAVVYKVRGP 172
Db 299 NEQDIPEGHDRTAQDFHEKTKNDWIYRKFRKMPGMSYVETDYSE-----FVGTNNGH 353

```

```

Qy 173 VRTVT--KRVQPS---LDPATOKLITNIFSKEMFNKTMALMDLDVKMPLGKLSKQOIA 227
Db 354 KKKITPGSKITPGSKITLLPKSVKEVVMISFDVENMKSAKSEIDVNMKPLGRLSHNQIN 413
Qy 228 RGFPALEALEAL-KGPTDGGQSLSELSHFYVYIPIHNFHSGSQPPPIINSPELLOAKKDM 286
Db 414 LAFEVLNDISDLLVKLPIDASKIL-DFSNGKFYTIIPHNFGRVPEPIDSFHKEKKNML 472
Qy 287 LVLADIELAQALQAVSEKQTVVEVPHPLDRDYQLKQQLQDLSGAPYKVIQTYLEQT 346
Db 473 NALLDIKFAYDQISGGDVPASTSLSIDPVDINRYRKLKCIIMEPLQOQCDNNMIHOYLKNT 532
Qy 347 -GSNHRCP-LOHIWKVNOEGEEDRFOAHSKLGNRKLWHGTNMAVAAILTSGLRIMPH 404
Db 533 HGATHDLKVELIDILKYNRDNESKFKRH--IGNRLLWHGSGKNMFAGILCOGLRIAPP 590
Qy 405 ----SGGRVKGIIYFASENSKSAGYVIGMKCGAHHVGYMFLGVALGR-EHHINTDNPSL 459
Db 591 EAPVSGYMGKGVYFADMFSKFFY---CRANAKEEAYLLLCDVALGNVQQLMASKNYSR 647
Qy 460 KSPPPGFDSTIARG-----HTEPD-----PTQDTELELDQVQVVPQCPPEFS 506
Db 648 QTLFAGFSQVQGVGRQCPREIGSKYKNDPGYTVPLGLTYMLOGKQNV----- 694
Qy 507 SSTFSQSEVLIYQESQCLRYLLEVHL 533
Db 695 DYHLLYNEFIVDQIQIQLKYLVRVKM 721

```

Search completed: August 29, 2002, 08:01:31  
Job time: 363 sec

**THIS PAGE BLANK (USPTO)**

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 07:57:14 ; Search time 124.84 seconds  
(without alignments)  
480.454 Million cell updates/sec

Title: US-09-701-586b-6  
Perfect score: 2854  
Sequence: 1 MSLLFLAMPKPKPWQTEG.....EVLIIQESQCLRLYLEVHL 540

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_032802.\*  
1: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2854	100.0	540	21 AAY51176	Human uterus type
2	2823	98.9	533	21 AAY51175	Human brain PARP3
3	2811	98.5	533	22 AAU29021	Human PARP-3 prote
4	2253.5	79.0	528	21 AAY51178	Murine PARP1 (shor
5	2241	78.5	533	21 AAY51177	Murine PARP1 (long
6	738	25.9	522	22 AAU29022	Mouse PARP-2 prote
7	729.5	25.6	583	22 AAB47029	hPARP2. Homo sapi
8	726.5	25.5	534	21 AAB42909	Human ORFX ORP2673
9	726.5	25.5	534	22 AAU29023	Human PARP-2 prote
10	725.5	25.4	534	22 AAU29020	Human PARP-2 prote
11	725	25.4	570	21 AAY51174	Human brain PARP2

12	725	25.4	570	22 AAB11480	Human brain poly-A
13	722	25.3	521	22 AAB60693	Human poly(ADP-rib
14	688.5	24.1	637	21 AAY68835	The poly(ADP-ribos
15	682.5	23.9	531	22 AAB93513	Human protein sequ
16	672	23.5	1014	21 AAY58043	Human poly (ADP-ri
17	672	23.5	1014	22 AAU29019	Human PARP-1 prote
18	672	23.5	1014	22 AAB66296	Human tankyrase2 r
19	670.5	23.5	653	21 AAY68834	A poly(ADP-ribose)
20	670.5	23.5	1014	21 AAY49939	Human nuclear NAD+
21	670	23.5	1013	17 AAY99642	Poly(ADP-ribose) p
22	664	23.3	1014	20 AAY33699	Human poly(ADP-rib
23	640.5	22.4	1063	22 AAB7032	Fusion protein PAR
24	638	22.4	557	22 ABB66431	Drosophila melanog
25	578	20.3	969	21 AAY68833	A poly(ADP-ribose)
26	578	20.3	980	21 AAY68839	A poly(ADP-ribose)
27	475	16.6	982	20 AAY28464	Maize poly ADP-rib
28	472.5	16.6	379	22 AAU21687	Novel human neopla
29	447	15.7	360	22 AAB47030	N-terminal fragmen
30	410.5	14.4	1010	21 AAY68840	Fusion protein of
31	394	13.8	294	22 AAU20129	Human DNA repair a
32	394	13.8	294	22 AAU21810	Novel human neopla
33	390	13.7	287	22 AAB47031	C-terminal fragmen
34	288	10.1	1099	22 AAB66301	Human tankyrase2 e
35	267.5	9.4	227	22 AAU20130	Human DNA repair a
36	267.5	9.4	227	22 AAU21811	Novel human neopla
37	258	9.0	1730	22 AAU33242	Novel human secret
38	256	9.0	1724	21 AAY34373	CDNA sequence enco
39	256	9.0	1724	22 AAB51022	Human minor vault
40	216.5	7.6	190	22 AAU21688	Novel human neopla
41	138.5	4.9	1327	21 AAB27212	Human tankyrase I
42	138.5	4.9	1327	21 AAY44402	Human tankyrase
43	138.5	4.9	1327	22 AAB66279	Human tankyrase1 S
44	137	4.8	756	22 AAB66286	Human tankyrase2 C
45	137	4.8	784	22 AAB66285	Human tankyrase2 C

ALIGNMENTS

RESULT 1  
AAY51176  
ID AAY51176 standard; Protein: 540 AA.  
XX  
AC AAY51176;  
XX  
DT 31-MAR-2000 (first entry)  
XX  
DE Human uterus type 2 PARP3 protein.  
DE  
DE PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP3.  
XX  
OS Homo sapiens.  
XX  
PN WO9964572-A2.  
XX  
PD 16-DEC-1999.  
XX  
PF 04-JUN-1999; 99WO-EP03889.  
XX  
PR 05-JUN-1998; 98DE-1025213.  
PR 01-MAR-1999; 99DE-1008837.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX WPI; 2000-087218/07.  
XX N-PSDB; AAZ44289.  
XX  
XX Novel genes and proteins, antibodies and binding partners useful in  
PT diagnosis and therapy of energy deficiency associated disease

PT conditions -  
XX Claim 4; Page 62-64; 96pp; German.  
XX  
XX This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>2C (I). The nucleic acid  
CC sequences, PARP homologues and antibodies are useful for analytic  
CC detection of PARP homologues and for identifying PARP effectors or  
CC binding partners, as well as for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
XX human PARP3 protein used in the method of the invention.  
XX  
SQ Sequence 540 AA;

Query Match 100.0%; Score 2854; DB 21; Length 540;  
Best Local Similarity 100.0%; Pred. No. 1.7e-252;  
Matches 540; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MSLLFLAMAPKPKPWVOTEGPEKKKGROAGREDDPFRSTAEALKAIPAERKRIIRVDPTCP 60  
DB 1 MSLLFLAMAPKPKPWVOTEGPEKKKGROAGREDDPFRSTAEALKAIPAERKRIIRVDPTCP 60  
QY 61 LSSNGPTQVYEDYNTLNQNIENNKKFYIIQLLODSNRRFFTCWNRWGRVGEVGSQKIN 120  
DB 61 LSSNGPTQVYEDYNTLNQNIENNKKFYIIQLLODSNRRFFTCWNRWGRVGEVGSQKIN 120  
QY 121 HFTRLDAKDFEKKFREKTKNNWAERDHFVSHPGKTYLIEVQADEAQAVKVDGRGPV 180  
DB 121 HFTRLDAKDFEKKFREKTKNNWAERDHFVSHPGKTYLIEVQADEAQAVKVDGRGPV 180  
QY 181 RTVTKRVQPSLDPATOKLTINIFSKEMFKNTMALMDLVKKMPLGLSKQQTARGFEAL 240  
DB 181 RTVTKRVQPSLDPATOKLTINIFSKEMFKNTMALMDLVKKMPLGLSKQQTARGFEAL 240  
QY 241 EALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPIINSPELLQAKKMDLLVLADIE 300  
DB 241 EALEEALKGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPIINSPELLQAKKMDLLVLADIE 300  
QY 301 LAQALQAVSEGEKTVVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTSNHRCP 360  
DB 301 LAQALQAVSEGEKTVVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTSNHRCP 360  
QY 361 TLOHIWKVNOGEGEDRFOAHKSLGNRKLWHGTNNMAVVAAILTSGLRIMPHSGSRVKGKI 420  
DB 361 TLOHIWKVNOGEGEDRFOAHKSLGNRKLWHGTNNMAVVAAILTSGLRIMPHSGSRVKGKI 420  
QY 421 YFASSENSKAGYVIGMKCGAHVGYMFLGEVALGREHHINTDNPSLKSPPPGSDVIARG 480  
DB 421 YFASSENSKAGYVIGMKCGAHVGYMFLGEVALGREHHINTDNPSLKSPPPGSDVIARG 480  
QY 481 HTEPDPTQDTELELDGQGVVVPQGPVPCPEFSSSTFSQSEYLIYQBSQRLRYLLEVHL 540  
DB 481 HTEPDPTQDTELELDGQGVVVPQGPVPCPEFSSSTFSQSEYLIYQBSQRLRYLLEVHL 540

RESULT 2  
AA51175  
ID AA51175 standard; Protein; 533 AA.  
AC  
AC AA51175;  
XX  
XX  
XX  
DT 31-MAR-2000 (first entry)  
XX  
DE Human brain PARP3 protein.  
XX

KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP3.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9964572-A2.  
PN  
XX  
XX 16-DEC-1999.  
PD  
XX  
XX 04-JUN-1999; 99WO-EP03889.  
PF  
XX  
XX 05-JUN-1998; 98DE-1025213.  
PR  
XX  
XX 01-MAR-1999; 99DE-1008837.  
PX  
XX  
XX (BADI ) BASF AG.  
PA  
XX  
XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
PI  
XX  
XX WPI: 2000-087218/07.  
DR  
XX  
XX N-PSDB; AA244288.  
DR  
XX  
XX Novel genes and proteins, antibodies and binding partners useful in  
PT diagnosis and therapy of energy deficiency associated disease  
PT conditions -  
PT  
XX  
XX Claim 4; Page 57-59; 96pp; German.  
PS  
XX  
XX This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>MX<sub>2</sub>2C (I). The nucleic acid  
CC sequences, PARP homologues and antibodies are useful for analytic  
CC detection of PARP homologues and for identifying PARP effectors or  
CC binding partners, as well as for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
CC human PARP3 protein used in the method of the invention.  
CC  
XX  
SQ Sequence 533 AA;

Query Match 98.9%; Score 2823; DB 21; Length 533;  
Best Local Similarity 100.0%; Pred. No. 1.1e-249;  
Matches 533; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 8 MAPKPKPWVOTEGPEKKKGROAGREDDPFRSTAEALKAIPAERKRIIRVDPTCLSSNPQT 67  
DB 1 MAPKPKPWVOTEGPEKKKGROAGREDDPFRSTAEALKAIPAERKRIIRVDPTCLSSNPQT 60  
QY 68 QVYEDYNTLNQNIENNKKFYIIQLLODSNRRFFTCWNRWGRVGEVGSQKINHTRLED 127  
DB 61 QVYEDYNTLNQNIENNKKFYIIQLLODSNRRFFTCWNRWGRVGEVGSQKINHTRLED 120  
QY 128 AKDDEKFKFREKTKNNWAERDHFVSHPGKTYLIEVQADEAQAVKVDGRGPVTVTKRV 187  
DB 121 AKDDEKFKFREKTKNNWAERDHFVSHPGKTYLIEVQADEAQAVKVDGRGPVTVTKRV 180  
QY 188 QPCSLDPATOKLTINIFSKEMFKNTMALMDLVKKMPLGLSKQQTARGFEALAEAL 247  
DB 181 QPCSLDPATOKLTINIFSKEMFKNTMALMDLVKKMPLGLSKQQTARGFEALAEAL 240  
QY 248 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPIINSPELLQAKKMDLLVLADIEALQALQA 307  
DB 241 KGPTDGGQSLEELSSHFYTVIPHNFGHSQPPPIINSPELLQAKKMDLLVLADIEALQALQA 300  
QY 308 VSEGEKTVVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTSNHRCPPTLOHIWK 367  
DB 301 VSEGEKTVVEVPHPLDRDYQLLKCQLQLLDSGAPEYKVIQTYLEQTSNHRCPPTLOHIWK 360

QY 368 VNOEGEEDRFOAHSKLGNRKLLWHGTNNVAAIITSLGRIMPHSGRGVKGIVFASENS 427  
Db 361 vnqgeeedrfqahsklgnrkllwhgtmmavvaalitsglrmpshggrvgkgyifasens 420  
QY 428 KSAGYVIGMKGAHHVGYMFLGEVALGREHINTDNPGLKSPPPGFDVSIARGHTEPDPT 487  
Db 421 ksagyvigmkgahhvgyymflgevalgrehntdnpslkspgpgfdsviarghtepdpt 480  
QY 488 QDTELELDGQOVVVPQGQVPCPEPSSSTFSQSEYLIYQESQCRLYRILEVHL 540  
Db 481 qdteleldgqgvvvpqgqvpvpcpefssstfsqseyliyqesqcrlyrillevhl 533

RESULT 3  
AAU29021  
ID AAU29021 standard; Protein: 533 AA.  
AC AAU29021;  
DT 18-DEC-2001 (first entry)  
DE Human PARP-3 protein.  
KW Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
KW cytosolic; nontropic; neuroprotective; antiinflammatory; antidiabetic;  
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
OS Homo sapiens.  
XX  
XX WO200164955-A1.  
XX  
XX 07-SEP-2001.  
XX  
XX 01-MAR-2001; 2001WO-US06572.  
XX  
XX 02-MAR-2000; 2000US-0517467.  
XX (ISIS-) ISIS PHARM INC.  
XX  
XX Popoff I, Cowser LM;  
XX  
XX WPI: 2001-602570/68.  
XX  
XX N-PSDB; AAS45590.  
XX  
XX Antisense compound useful for treating hyperproliferative,  
XX neurological, inflammatory and autoimmune disorders and diabetes  
XX inhibits human PARP -  
XX  
XX Example 13; Page 105-107; 168pp; English.  
XX  
XX The invention relates to antisense oligonucleotides targeted to human  
XX PARP nucleic acid and inhibiting expression of human PARP. PARP  
XX (Poly (ADP-ribose) polymerase plays an important role in chromatin  
XX decondensation, DNA replication, DNA repair, gene expression, malignant  
XX transformation, cellular differentiation and apoptosis. The antisense  
XX oligonucleotide inhibitors are useful for inhibiting the expression of  
XX PARP in human cells or tissues. They are also useful for treating a  
XX human with a disease associated with PARP especially hyperproliferative  
XX disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
XX neurological (e.g. parkinsonism, meningitis-associated intracranial  
XX complications and ischaemia), inflammatory and autoimmune disorders (e.g  
XX arthritis) and diabetes. The present sequence is a PARP protein,  
XX the cDNA encoding which was used to design the antisense  
XX oligonucleotides.  
XX  
XX Sequence 533 AA;

Query Match 98.5%; Score 2811; DB 22; Length 533;

Best Local Similarity 99.6%; Pred. No. 1.4e-248;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 MAPKPKVWQTEGPEKKKGROAGREEDPFRSTAALKAIAPAEKRIIRVDPPTCPCLSSNPGT 67  
Db 1 mapkpkvwqtegpekkkgroagreedpfrstaalkaipaeakriirvdpctplssnpgt 60  
QY 68 QVYEDYNCTLNQNTENNKNFYIQLLODSNRFTCNRRGRCVGEVGSINHTFRLD 127  
Db 61 qvyedynctlnqntennknkfyliqlldsnrftcnrrgrvgevgslnhfrld 120  
QY 128 AKKDFEKKFKRETKNNWAERDHFVSHPGKYTLIEVQAEDAEQAVVWVDRGPVTRVTKRV 187  
Db 121 akkdfekfkrektnnwaerdhfvshpgkytlievqaedaeavvkvdrapvtrvtrv 180  
QY 188 QPCSLDPATQKLIITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQIARGFFALEALEAL 247  
Db 181 qpcslldpatqkliitnifskemfkntmalmdldvkkmplgklskqiarqgfealealeal 240  
QY 248 KGPTDGGQSLSELSHFYTVIPHNFGHSQPPPINSPELLQAKKDMLLVLADIELAQA 307  
Db 241 kgptdggqslselelshfycviphnfgshsqpppinnspellqakkdmllvadielaqala 300  
QY 308 VSEKTEVEEYVPHPLDRDYQLLKQQLDSCAPEYKVIQTYLEQTSNHRCPQLQHIWK 367  
Db 301 vsekteveeyvphpldrdyqlllkqqlldsgapeykvityleqtsnhrcptlqhiwk 360  
QY 368 VNQGEEDRFOAHSKLGNRKLLWHGTNNVAAIITSLGRIMPHSGRGVKGIVFASENS 427  
Db 361 vnqgeeedrfqahsklgnrkllwhgtmmavvaalitsglrmpshggrvgkgyifasens 420  
QY 428 KSAGYVIGMKGAHHVGYMFLGEVALGREHINTDNPGLKSPPPGFDVSIARGHTEPDPT 487  
Db 421 ksagyvigmkgahhvgyymflgevalgrehntdnpslkspgpgfdsviarghtepdpt 480  
QY 488 QDTELELDGQOVVVPQGQVPCPEPSSSTFSQSEYLIYQESQCRLYRILEVHL 540  
Db 481 qdteleldgqgvvvpqgqvpvpcpefssstfsqseyliyqesqcrlyrillevhl 533

## RESULT 4

AAU51178  
ID AAU51178 standard; Protein: 528 AA.  
AC AAU51178;  
XX  
XX 31-MAR-2000 (first entry)  
XX  
XX Murine PARP1 (short) homologue protein.  
XX  
XX PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
XX diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
XX ischemic tissue damage; PARP1.  
XX  
XX Mus sp.  
XX  
XX WO9964572-A2.  
XX  
XX 16-DEC-1999.  
XX  
XX 04-JUN-1999; 99WO-EP03889.  
XX  
XX 05-JUN-1998; 98DE-1025213.  
XX  
XX 01-MAR-1999; 99DE-1008837.  
XX  
XX (BADI ) BASF AG.  
XX  
XX Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;  
XX  
XX WPI: 2000-087218/07.  
XX  
XX N-PSDB; AAZ44291.  
XX  
XX Novel genes and proteins, antibodies and binding partners useful in  
XX  
XX PT

PT diagnosis and therapy of energy deficiency associated disease  
PT conditions -

XX  
XX  
PS Claim 4; Page 71-73; 96pp; German.

XX This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>2MHX<sub>2</sub>2C (I). The nucleic acid  
CC detection of PARP homologues and antibodies are useful for analytic  
CC PARP-binding partners are useful for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
CC murine PARP1 protein used in the method of the invention.

XX Sequence 528 AA;

Query Match 79.0%; Score 2253.5; DB 21; Length 528;  
Best Local Similarity 80.9%; Pred. No. 1.8e-197;  
Matches 432; Conservative 37; Mismatches 58; Indels 7; Gaps 4;

QY 8 MAPKPKPWQTEGPEKKGRQAGREEDPRSTAEALKAIPAERKRIIRVDPCTCLSSNPCT 67  
DB 1 mapkrkasvqtg--skkqrqgteedsfrstaealraapadnrvirvdpscfnsrpgi 58  
QY 68 QYVEDYNTLQNTNIENNKKFYIIQLQDSNRFCTCWRNWRGVGSGKINHFTRLED 127  
DB 59 qvhedydctlnqtnignnnkfyilqlleegsrff-cwnrgrvrgvgqskmnhfctled 117  
QY 128 AKKDFEKKFKRKTNNWAERDHFVSHPGKYTLIEVQAEDEAQAQVAVKVRGPTTKRV 187  
DB 118 akkdfkkfkwekcknwkeerdfrvaqpnkytlievgeaesgeavvkdsgpvrty---v 174  
QY 188 QPCSLDPATQKLTNIFSKEMFKNTMALMDLVKKMPLGKLSKQOITARGFEALEEAL 247  
DB 175 kpcslpatqnlitnifskemfknamtlmndvkkmpkgkitkqkiargfealeeaa 234  
QY 248 KGPTDGGQSLSEELSSHYTVIPHNFGHSQPPPINSPPELLQAKKMDLLVLADIQAALQA 307  
DB 235 knptdggslseelsfcyfviphnfgrspppnsdvqlqakdmllvdiadlaqlqa 294  
QY 308 V-SEQEKTVEEVPHPDLDRDYQLLKCQLQLDSCGAPEYKVIQTLYEQTGNSNHRCPPTQHIW 366  
DB 295 apgeeeekveevphidrdyqlircqlqldsgeseykaigtlykqtgnsyrcpnlrhv 354  
QY 367 KVNQEGEDRFOAHSKLGNRKLLWHGTNMAVAAITSLGRIMPHSGRGVKGIGYFASEN 426  
DB 355 kvnregegdfrqahsklgnrllwhgtcnvavvaaitslgrlmphsggrvgkigyfasen 414  
QY 427 SKSAGYVIGMKCAHGYMFLGEALGRBHHINTDNPSLSKSPPPGSDVSFARGHTEPDP 486  
DB 415 sksagvyttmhcgghqvgymflgevalgkehhtidpkslkpppgfsviargtpepd 474  
QY 487 QTDELELDGQVVPVQGVPCPEPSSSTFSQSEYLIYQESQRLRYLLEVHL 540  
DB 475 aqdieleldgqvvpvpgppvpqcpstfssfsdseyliysesqcrilrylleihl 528

RESULT 5

AA51177  
ID AA51177 standard; Protein; 533 AA.

XX  
AC AA51177;

XX  
DT 31-MAR-2000 (first entry)

DE Murine PARP1 (long) homologue protein.

XX PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;  
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;  
KW ischemic tissue damage; PARP1.

OS Mus sp.

XX WO9964572-A2.

XX 16-DEC-1999.

XX 04-JUN-1999; 99WO-EP03889.

XX 05-JUN-1998; 98DE-1025213.

PR 01-MAR-1999; 99DE-1008837.

XX (BADI ) BASF AG.

PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;

XX WPI: 2000-087218/07.

DR N-PSDB; AA244290.

XX Novel genes and proteins, antibodies and binding partners useful in  
PT diagnosis and therapy of energy deficiency associated disease  
PT conditions -

XX Claim 4; Page 67-69; 96pp; German.

PS This invention describes novel human and murine poly(ADP-ribose)  
CC polymerase (PARP) homologues, which are characterised by an amino acid  
CC sequence with a functional NAD<sup>+</sup>-binding site and no zinc finger  
CC sequence motif, of general formula CX<sub>2</sub>CX<sub>2</sub>2MHX<sub>2</sub>2C (I). The nucleic acid  
CC sequences, PARP homologues and antibodies are useful for analytic  
CC detection of PARP homologues and for identifying PARP effectors or  
CC binding partners, as well as for determining their effectiveness.  
CC PARP-binding partners are useful for the diagnosis or therapy of a  
CC disease condition, which is the result of a PARP protein, especially an  
CC energy deficiency, which may comprise tissue damage from cell death  
CC following necrosis or apoptosis. The disease condition may be chosen  
CC from a neurodegenerative illness, or sepsis or ischemic tissue damage,  
CC in particular neurotoxic disturbances, etc. This sequence represents the  
CC murine PARP1 protein used in the method of the invention.

XX Sequence 533 AA;

Query Match 78.5%; Score 2241; DB 21; Length 533;  
Best Local Similarity 80.1%; Pred. No. 2.5e-196;  
Matches 432; Conservative 37; Mismatches 58; Indels 12; Gaps 5;

QY 8 MAPKPKPWQTEGPEKKGRQAGREEDPRSTAEALKAIPAERKRIIRVDPCTCLSSNPCT 67

DB 1 mapkrkasvqtg--skkqrqgteedsfrstaealraapadnrvirvdpscfnsrpgi 58

QY 68 QYVEDYNTLQNTNIENNKKFYIIQLQDSNRFCTCWRNWRGVGSGKINHFTRLED 127

DB 59 qvhedydctlnqtnignnnkfyilqlleegsrff-cwnrgrvrgvgqskmnhfctled 117

QY 128 AKKDFEKKFKRKTNNWAERDHFVSHPGKYTLIEVQAEDEAQAQVAVK----VDRGPVRT 182

DB 118 akkdfkkfkwekcknwkeerdfrvaqpnkytlievgeaesgeavvkdsgpvrty 177

QY 183 VTKRVPCLSDPATQKLTNIFSKEMFKNTMALMDLVKKMPLGKLSKQOITARGFEALEA 242

DB 178 v---vkpcslpatqnlitnifskemfknamtlmndvkkmpkgkitkqkiargfealea 234

QY 243 LEALKGPTDGGQSLSEELSSHYTVIPHNFGHSQPPPINSPPELLQAKKMDLLVLADIELA 302

DB 235 leeamkntpgdggslseelsfcyfviphnfgrspppnsdvqlqakdmllvdiadla 294

QY 303 QALQAV-SEQEKTVEEVPHPDLDRDYQLLKCQLQLDSCGAPEYKVIQTLYEQTGNSNHRCP 361

Db 295 qtlqaapgeeeekveevphpldrdyqllrqqlldgeseykaiqylkqgtgnsyrcpn 354  
 QY 362 LOHTWKYNOEGEEDRFOAHSGKLNKRLWHTGNMAVVAAILTSGLRMPHSGGRVGKGIY 421  
 Db 355 lrhvwknvregedrfqahsklgnrrllwghtnavvaailtsglrmpghsggrvkgiy 414  
 QY 422 FASENSKSAGYVIGMKGAHVGYMFLGEVALGREHINTDNPSLKSPPPGFDSVTARGH 481  
 Db 415 fasensksagvyvtmhcgghqvgymfigevalgkehhitiddpslkpppgfsgviargq 474  
 QY 482 TEPDPTQDTELELDGQQVVPVQGPVPCPEFSSSTFQSEYLYQESQRLRYLLEVHL 540  
 Db 475 tepdaqdieleldgqvvpvpgpvcpsfkssfsqseylykesqcrlylleihl 533

RESULT 6  
 AAU29022  
 ID AAU29022 standard; Protein; 522 AA.  
 XX AC  
 AAU29022;  
 XX DT  
 XX 18-DEC-2001 (first entry)  
 XX DE  
 DE Mouse PARP-2 protein.  
 XX Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
 KW cytosolic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;  
 KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
 KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
 KW meningitis-associated intracranial complication; ischaemia;  
 KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
 XX OS  
 XX Mus musculus.  
 XX WO200164955-A1.  
 XX PN  
 XX 07-SEP-2001.  
 XX 01-MAR-2001; 2001WO-US06572.  
 XX 02-MAR-2000; 2000US-0517467.  
 XX (ISIS-) ISIS PHARM INC.  
 XX Popoff I, Cowsett LM;  
 XX WPI; 2001-602570/68.  
 XX N-PSDB; AAS45597.  
 XX Antisense compound useful for treating hyperproliferative,  
 PT neurological, inflammatory and autoimmune disorders and diabetes  
 PT inhibits human PARP -  
 XX  
 PS Example 13; Page 109-111; 168pp; English.  
 XX The invention relates to antisense oligonucleotides targeted to human  
 CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
 CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
 CC decondensation, DNA replication, DNA repair, gene expression, malignant  
 CC transformation, cellular differentiation and apoptosis. The antisense  
 CC oligonucleotide inhibitors are useful for inhibiting the expression of  
 CC PARP in human cells or tissues. They are also useful for treating a  
 CC human with a disease associated with PARP especially hyperproliferative  
 CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
 CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
 CC complications and ischaemia), inflammatory and autoimmune disorders (e.g  
 CC arthritis) and diabetes. The present sequence is a PARP protein,  
 CC the cDNA encoding which was used to design the antisense  
 CC oligonucleotides.  
 XX SQ  
 Sequence 522 AA;

Query Match 25.9%; Score 738; DB 22; Length 522;  
 Best Local Similarity 34.3%; Pred. No. 1.5e-58;  
 Matches 187; Conservative 103; Mismatches 191; Indels 64; Gaps 19;

QY 23 KKGRQA-GREEDPFRSTABALKAIPAEKRIIRVDPTCLSSNPGTOVY----EDYNCTL 77  
 Db 5 qrkqpmaggkdadrtdkndrsvktlllkkgk-avdpcecaaklkg-ahvycegddvydml 62  
 QY 78 NOTNIENNKKFYIQLLO--DSNRFETCWNRWGRVGEVGO-SKINHFTRLEDAKKDFEKK 135  
 Db 63 nqnllqfnnnkylqlldeddagrnsvmwrgrvgtkgqnslvctsgdlnkakeifqkk 122  
 QY 136 FREKTKNNWAERDFHVSHPGKYTLIEV---QAEDEA---QEAAYKVDGRGPVTRVRVQ 188  
 Db 123 fldtknnwedrenfekvpgkymldmqdaastqdesktkeetlkpe----- 170  
 QY 189 PCSLDPATQKLTINIFSKEMFKMTMALMDLDVKKMPLGKLSKQOIARGFPALEALEALK 248  
 Db 171 -sgldlrvgellklcnvqmeemmlenkydtkraplgtvqakgagyslkkiedcir 229  
 QY 249 GPTDGGOSLELSHFYVIPHFGHSOPPPINSPELLOAKKMLLVLADELAQAALQAV 308  
 Db 230 a-qghgralveacnefyiripdhfgisipvirtekeiskvkllealgdlei--alkv 286  
 QY 309 SEQKTVVEEYVPHPLDRDYQLLKQQLLDGSAPEYKVIQTYLEQT-GSNHR--CPTLQHI 365  
 Db 287 kserqgle--hpldqhyrnlhcalrpldhesnefkvisqylqsthapthkhtytildv 343  
 QY 366 WKYNQEGEEDRFOAHSGKLNKRLWHTGNMAVVAAILTSGLRIMPH----SGGRVKGKIY 421  
 Db 344 fevekegekeaf--edlpnrmllwhgsrlsnvvgilshglrvappeapitgymfgkgy 401  
 QY 422 FASENSKSAGYVIGMKGAHVGYMFLGEVALGREHINTDNPSLKSPPPGFDSVIARGH 481  
 Db 402 fadmsksanycfasr--lknrtgillisevalgqcnleeanpkagqllrghsktgmgk 459  
 QY 482 TEPDPTQDTELELDGQQVVPVQGPVPCPEFSSS-----TFQSEYLYQESQRLR 533  
 Db 460 mapsa-----hftlmgstvpplgpasdtgilnpegyltnyefivyspnqvr 509  
 QY 534 YLLEV 538  
 Db 510 YLLKI 514

RESULT 7  
 AAB47029  
 ID AAB47029 standard; Protein; 583 AA.  
 XX AC  
 XX AAB47029;  
 XX DT  
 XX 29-MAR-2001 (first entry)  
 XX DE  
 XX hPARP2.  
 XX Human; poly(ADP-ribose) polymerase; hPARP2; oxidative stress; ARDS;  
 KW inflammation; ischaemic stroke; hemorrhagic shock; myocardial ischemia;  
 KW infarction; cerebral vasospasm; rheumatoid arthritis; osteoarthritis;  
 KW gouty arthritis; spondylitis; Behcet's disease; sepsis; septic shock;  
 KW endotoxic shock; gram negative sepsis; gram positive sepsis; trauma;  
 KW toxic shock syndrome; multiple organ injury syndrome; vasculitis;  
 KW hemorrhage; conjunctivitis; uveitis; thyroid-associated ophthalmopathy;  
 KW eosinophilic granuloma; asthma; chronic bronchitis; allergic rhinitis;  
 KW chronic obstructive pulmonary disease; silicosis; reperfusion injury;  
 KW pulmonary sarcoidosis; pleurisy; alveolitis; pneumonia; myocardium;  
 KW bronchiectasis; pulmonary oxygen toxicity; keloid formation; brain;  
 KW scar tissue formation; atherosclerosis; systemic lupus erythematosus;  
 KW autoimmune thyroiditis; multiple sclerosis; Reynaud's syndrome;  
 KW graft versus host disease; allograft rejection; cystic fibrosis;  
 KW chronic glomerulonephritis; inflammatory bowel disease;  
 KW Crohn's disease; ulcerative colitis; necrotizing enterocolitis;  
 KW inflammatory dermatosis; contact dermatitis; atopic dermatitis;  
 KW psoriasis; urticaria; fever; myalgia; meningitis; encephalitis;

KW Sjogren's syndrome; alcoholic hepatitis; bacterial pneumonia;  
 KW hypovolemic shock; type 1 diabetes mellitus; hypersensitivity;  
 KW leukocyte dyscrasia; thermal injury; cytokine-induced toxicity.  
 XX  
 OS Homo sapiens.  
 XX WO200077179-A2.  
 XX  
 XX  
 PD 21-DEC-2000.  
 XX  
 PF 16-JUN-2000; 2000WO-US16629.  
 XX  
 XX 16-JUN-1999; 99US-0139543.  
 PR  
 XX (ICOS-) ICOS CORP.  
 XX  
 XX Christenson E, Demaggio AJ, Goldman PS, McElligott DL;  
 XX WPI; 2001-025335/03.  
 DR N-PSDB; AAC85303.  
 XX  
 XX New human poly(ADP-ribose) polymerase for treating inflammatory,  
 PT neurological, cardiovascular, or neoplastic tissue growth disorders,  
 PT such as, arthritis, encephalitis, myocardial ischemia, and leukocyte  
 PT metastasis.  
 XX  
 XX Claim 3; Page 94-95; 129pp; English.  
 XX  
 XX This sequence represents human poly(ADP-ribose) polymerase (hPARP2).  
 CC This protein causes the covalent addition of polymers of ADP-ribose  
 CC to protein targets. hPARP2 activity is induced in many instances of  
 CC oxidative stress or during inflammation where there is direct damage  
 CC to the DNA. hPARP2 may be used to identify antagonists which  
 CC may be used to treat a human having a disorder mediated by PARP2  
 CC activity, such as, inflammatory, neurological, cardiovascular,  
 CC or neoplastic tissue growth disorders, e.g. ischemic stroke,  
 CC hemorrhagic shock, myocardial ischemia or infarction,  
 CC transplantation, cerebral vasospasm; rheumatoid, osteo- or gouty  
 CC arthritis, spondylitis; Behcet's disease; sepsis, septic or endotoxic  
 CC shock, gram negative or positive sepsis, toxic shock syndrome; multiple  
 CC organ injury syndrome secondary to septicemia, trauma, or hemorrhage;  
 CC allergic or vernal conjunctivitis, uveitis, thyroid-associated  
 CC ophthalmopathy; eosinophilic granuloma; asthma, chronic bronchitis,  
 CC allergic rhinitis, ARDS, chronic obstructive pulmonary disease,  
 CC silicosis, pulmonary sarcoidosis, pleurisy, alveolitis, vasculitis,  
 CC pneumonia, bronchiectasis, pulmonary oxygen toxicity; reperfusion  
 CC injury of the myocardium, brain or extremities; cystic fibrosis; keloid  
 CC formation, scar tissue formation; atherosclerosis; systemic lupus  
 CC erythematosus, autoimmune thyroiditis, multiple sclerosis; Reynaud's  
 CC syndrome; graft versus host disease, allograft rejection; chronic  
 CC glomerulonephritis; inflammatory bowel disease, Crohn's disease,  
 CC ulcerative colitis, necrotizing enterocolitis; inflammatory dermatoses,  
 CC contact or atopic dermatitis, psoriasis, urticaria, fever and myalgias  
 CC due to infection; meningitis, encephalitis, and brain and spinal cord  
 CC injury due to minor trauma; Sjogren's syndrome; diseases involving  
 CC leukocyte diapedesis; alcoholic hepatitis; bacterial pneumonia;  
 CC antigen-antibody complex mediated diseases; hypovolemic shock; Type 1  
 CC diabetes mellitus; acute and delayed hypersensitivity; disease states  
 CC due to leukocyte dyscrasia and metastasis; thermal injury; granulocyte  
 CC transfusion associated syndromes; and cytokine-induced toxicity.  
 CC hPARP2 and antibodies to it, can also be used to diagnose these  
 CC conditions.  
 XX  
 XX Sequence 583 AA;

Query Match 25.6%; Score 729.5; DB 22; Length 583;  
 Best Local Similarity 35.4%; Pred. NO. 1.1e-57;  
 Matches 201; Conservative 95; Mismatches 184; Indels 87; Gaps 22;  
 QY 22 EKKKGKQAG-----REED-----PFRSTA-----EALKAIPAERKRIIVDPCTPLSSNP 65  
 DB 46, eskkmpvaggkankrdtkdqgmpgrsvskrvskvlllkgk-apvdpcc--takv 102

QY 66 G-TQVY----EDYNCTLTNOTNIENNNNKPYITQLQ-DSNREFTCWNRMGRVGEVQ-SK 118  
 DB 103 gkahycegdvymlnqtnlqfnmkyliqllddaqrnfsvmrgvrvgmgqshl 162  
 QY 119 INHTRLEDAKKDFKFKREKTKNNMAERHDFVSHPGKYTLIEV----QAEDEAQEAUVK 174  
 DB 163 vacsgnlmkakeifqkklfdtknnwedrekfekypgykdmqlmqdyatntqdeetkkee 222  
 QY 175 VDRGPVRTVKRVQPCSLDPATQKLITNIFSKEMKNTMALMDLDVKKMPGKLSKQQA 234  
 DB 223 slksplkpesq-----ldlrvgelikicnvqameemmmemkntkkaplgklvtvaik 276  
 QY 235 RGFEALEALEALKGPTDGGQSLELSHFYTVIPHNFGHSQPPPTNSPELQAKKMDLL 294  
 DB 277 ayyqskkkiedcira-qghgralmecacnefytriphdfglrtpplirtqkelskqliie 335  
 QY 295 VLADIELAOLAOVSEQKTVVEVP-HPIDRDYQLLKCOLQLDLSGAPYKYVQTYLEQT 353  
 DB 336 algdieiaikl-----vktelqspehpldqhyrnlhcalrpldhesyefkvisgylqst 389  
 QY 354 GS---NHRCPPTLOHTWKVYNQEGEEDRFQAHSKLGNRKLHLWHGFTNMAVVAAILTSLRIMP 410  
 DB 390 haphsdytmildifevkdgekafr--edlhnrmllwhgsmnswwgllshglriap 447  
 QY 411 H----SGGRVKGKIYFASSENSKSGYVIGMKCGAHVGYMFLGEVALGHEHREHINTDNPSL 466  
 DB 448 peapitgymfgkiyfadmsksanycfasr--lknrtgllllsevalgcnelleanpka 505  
 QY 467 -----KSPPPGFSVITARGHTEP-DPTQDTELELDGQVVVPGQVPVCPDE 511  
 DB 506 egllgqkhtskglgmapssahfvltngstvpigpasdt-----gilnpdgy----- 552  
 QY 512 FSSSFQSEXYLIQESQCLRYLLEV 538  
 DB 553 ----clnyneyivnpnqvmryllkv 575  
 RESULT 8  
 AAB42909  
 ID AAB42909 standard; Protein; 534 AA.  
 XX AAB42909;  
 AC  
 DT  
 DE  
 DE Human ORFX ORF2673 polypeptide sequence SEQ ID NO:5346.  
 XX  
 KW Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;  
 KW vulnary; antipsoriatic; antiparkinsonian; nootropic; neuroprotective;  
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;  
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;  
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;  
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;  
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;  
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;  
 KW cardiovascular disease; diabetes mellitus; erythematoidism; SCID; AIDS;  
 KW cholesterol ester storage; systemic lupus erythematosus; infection;  
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;  
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;  
 KW bone damage; cartilage damage; antinflammatory disease; coagulation;  
 KW thrombosis; contraceptive.  
 OS Homo sapiens.  
 XX  
 PN WO200058473-A2.  
 XX  
 PD 05-OCT-2000.  
 XX  
 XX 31-MAR-2000; 2000WO-US08621.  
 XX  
 PR 31-MAR-1999; 99US-0127607.  
 PR 02-APR-1999; 99US-0127636.



PR 05-APR-1999; 99US-0127728.  
PR 30-MAR-2000; 2000US-0540763.  
PA (CURA-) CURAGEN CORP.  
XX  
XX Shinkets RA, Leach M;  
PI WPI; 2000-602362/57.  
XX N-PSDB; AAC77118.  
XX  
XX Novel nucleic acids and peptides derived from open reading frame X,  
PT useful for treating e.g. cancers, proliferative disorders,  
PT neurodegenerative disorders and cardiovascular disease -  
XX  
XX Claim 11; Page 4522-4524; 5507pp; English.  
XX  
XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,  
CC which represent the human ORF open reading frames 1 to 3161. The ORF  
CC sequences have activities such as: cytostatic; hepatotropic; vulnerary;  
CC antiparasitic; antiparkinsonian; neurotropic; neuroprotective;  
CC osteopathic; anticonvulsant; antiarthritic; immunosuppressant;  
CC immunostimulant; cardiant; thrombolytic; coagulant; vasotropic;  
CC antidiabetic; hypotensive; dermatological; immunosuppressive;  
CC antiinflammatory; antibacterial; antiviral; antifungal; antirheumatic;  
CC antithyroid; and antianaemic. The sequences can be used for determining  
CC the presence of or predisposition to, or preventing or treating  
CC pathological conditions associated with an ORF-associated disorder. The  
CC nucleic acids can be used to express ORF proteins in gene therapy  
CC vectors. The proteins and nucleic acids may be used to treat cancers,  
CC proliferative disorders, neurodegenerative disorders, osteoarthritis,  
CC graft vs host disease, cardiovascular disease, diabetes mellitus,  
CC hypertension, hypothyroidism, cholesterol ester storage, systemic lupus  
CC erythematosus, severe combined immunodeficiency (SCID), AIDS, viral,  
CC bacterial or fungal infection, malaria, autoimmune disorders, asthma,  
CC allergies, aplastic anaemia, burns, wounds, bone and cartilage damage,  
CC nocturnal haemoglobinuria, antiinflammatory disease; to enhance  
CC coagulation; to inhibit thrombosis; and as a contraceptive.  
XX  
XX Sequence 534 AA;

Query Match 25.5%; Score 726.5; DB 21; Length 534;  
Best Local Similarity 35.1%; Pred. No. 1.8e-57;  
Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;  
QY 20 GPEKKKGQAGREED-PPRSTA-----EALKAIPAERKRIIRVDPTCLSSNPG-TQVY-- 70  
D 5 gggkankdrtdkqdgmpgrswskrvsesvkalllkgk-avdpdec--takvgkahvyce 61  
QY 71 --EDYNTLNTNTENNKNKFIYIQLLO-DSNFFTCNWRGRVGEVGO-SKINHFTRELE 126  
D 62 gndvydmlnqnlqfnknkylqlldedqgrnfrvwmrgvrgkmqghslvacsgnln 121  
QY 127 DAKKDFEKKFREKTNNWAERDFVSHPGKYTLIEV-----QAEDEAQBAVVKVDRGPVRT 182  
D 122 kakelfgkfdkdknwnwedrekfekyvgkydmlqmdyatntqdeetkteeslkspklp 181  
QY 183 VTKRVQCSLDPAQKLTNTIFSEKFNKNTWALMDLVKMPGLSKSQQTARFEALEA 242  
D 182 esq-----ldrvqelkiklcnvgameememkntkkaplglktvaqikagvyslkk 235  
QY 243 LEEALKGPTDGGQSELESFHYVIPHNFHGSOPPPINSPQLQAKMDLLVLADIELA 302  
D 236 iedcira-qghralmeacnefytriphdfgrtpplrtrkqekselqlllealqldleia 294  
QY 303 QALQAVSEKTYEEVP-HPLDRDYQLKCOLQLLDSGAPEKYIQTLYEQTGS---NHR 358  
D 295 ikl-----vktelqsphepldqhrynlhcalrpldhesyefkvisgylqsthapthdsy 348  
QY 359 CPTLQHLWKNVQEGEEDRFQAHSLKGRKLLWHGHTNANVAAIITSLGRIMPH-----SGG 414  
D 349 tmtldlfevkdgekafr--edlhnmllwhgsrmsnwvgilshglrlappeatitgy 406

QY 415 RVKGIYFASENSKSAGYVIGMKGAHVGMFLGEVALGREHHTINTDNPSL----- 466  
D 407 mfgkgyifadmsksksanycfsar--lknktgilllsevalgcnelleanpkagllqgkh 464  
QY 467 -----KSPPGPFDSVIARGHTEP-DPTQDTELELDGGOVVVPOGPVPCPEFSSTFSQ 519  
D 465 stkglgkmapssahftlngstvpigpasdt-----gilnpdgy-----tlny 507  
QY 520 SEYLIYQESQCLRLRYLLEV 538  
D 508 neiyvnpnqvmryllkv 526  
RESULT 9  
AAU29023  
ID AAU29023 standard; Protein; 534 AA.  
XX  
XX AAU29023;  
XX AC  
XX 18-DEC-2001 (first entry)  
XX DE Human PARP-2 protein #2.  
XX  
XX Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;  
KW cytostatic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;  
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;  
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;  
KW meningitis-associated intracranial complication; ischaemia;  
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.  
XX  
XX Homo sapiens.  
XX WO200164955-A1.  
XX 07-SEP-2001.  
XX 01-MAR-2001; 2001WO-US06572.  
XX 02-MAR-2000; 2000US-0517467.  
XX (ISIS-) ISIS PHARM INC.  
XX Popoff I, Cowser LM;  
XX WPI; 2001-602570/68.  
XX N-PSDB; AAS45684.  
PT Antisense compound useful for treating hyperproliferative,  
PT neurological, inflammatory and autoimmune disorders and diabetes  
PT inhibits human PARP -  
XX Example 16; Page 125-127; 169pp; English.  
XX  
XX The invention relates to antisense oligonucleotides targeted to human  
CC PARP nucleic acid and inhibiting expression of human PARP. PARP  
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin  
CC decondensation, DNA replication, DNA repair, gene expression, malignant  
CC transformation, cellular differentiation and apoptosis. The antisense  
CC oligonucleotide inhibitors are useful for inhibiting the expression of  
CC PARP in human cells or tissues. They are also useful for treating a  
CC human with a disease associated with PARP especially hyperproliferative  
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,  
CC neurological (e.g. parkinsonism, meningitis-associated intracranial  
CC complications and ischaemia), inflammatory and autoimmune disorders (e.g.  
CC arthritis) and diabetes. The present sequence is a PARP protein,  
CC the cDNA encoding which was used to design the antisense  
XX oligonucleotides.  
XX  
XX Sequence 534 AA;

Query Match 25.5%; Score 726.5; DB 22; Length 534;  
Best Local Similarity 35.1%; Pred. No. 1.8e-57;

```
Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;
QY 20 GPEKKKGROAGREED--PFRSTA-----EALKAIPAKRIIRVDPCTPLSSNPG--TOVY-- 70
Db 5 gkankdrtdkdgmpgrswaskrvsesvkalllkgk-avdpdec--takvgkahvyce 61
QY 71 --BDYNTLNTQNTNIENNKKFYIIQLLO-DSNRFFTCWNRWGRVGEVGO--SKINHFTRIE 126
Db 62 gndvydvmnqnlqfnnnkyyliqlleddaqrnfswmrwgrvkgmghslvacsgnln 121
QY 127 DAKKDFEKKFREKTKNNAERDFHVSHPGKYTLIEV-----QAEDAEQAEAVVKVDRGPVRT 182
Db 122 kakeifgkfldtknwnedrekfepvgkydmlgmdayatntqdeetkkeslksplkp 181
QY 183 VTKRVQPCSLDPATQKITNIFSKEMFKNTMALMDLDVKKMPGLKSLKQOIAFGFEALEA 242
Db 182 esq-----ldlrqvclikicnvqameemmemknttkaplgkltvaqikagysgslk 235
QY 243 LEBALKGPTDGGOSLELSHFYTVIPHNFHGSQPPPIINSPELLQAKKMDLLVLADIELA 302
Db 236 iedcira-gqhgralmeacnefytriphdfglrtptlirtqkelsekiqilealadiea 294
QY 303 QALQAVSEQEKTVVEVP--HPLDRDYQLLKCQLLQSLDGAPEYKVIQTYLEQTGS----NHR 358
Db 295 ik1-----vktelqspenpldghyrlhcalrpldhesyefkvisqyiqsthapthsd 348
QY 359 CPTLQHIWKVNOEGEDRFOAHSLGNRKLLWHGTNNVAAIILTSGLRIMPH-----SGG 414
Db 349 tmlldlfevekdgekeaf--edlhnrmllwhgstrmsnvwgilshglrliapeapitgy 406
QY 415 RVGKGIYFASSENSKAGYVIGMKGCAHHVGYMFLGEVALGREHHINTDNPSSL----- 466
Db 407 mfgkglyfadmsksanycfasi--lkntglllvalgcnelleanpkagellqgkh 464
QY 467 -----KSPPPGDSVIARGHTPEP-DPTQDTELELQDQVVPVQGPVPCPEPSSSFFSQ 519
Db 465 stkglgkmapssahftlngstvtplgpasdt-----gilnpgdy-----tlny 507
QY 520 SEYLIYQESOCRLRYLLEV 538
Db 508 neyivnpgvrmryllkv 526

RESULT 10
AAU29020
ID AAU29020 standard; Protein; 534 AA.
AC AAU29020;
XX
XX 18-DEC-2001 (first entry)
DT
DE Human PARP-2 protein #1.
XX
XX Human; PARP; Poly (ADP-ribose) polymerase; antisense oligonucleotide;
KW cytosolic; neurotropic; neuroprotective; antiinflammatory; antidiabetic;
KW immunosuppressant; hyperproliferative disorder; cancer; cellular injury;
KW oxidative stress; neurological disorder; parkinsonism; apoptosis;
KW meningitis-associated intracranial complication; ischaemia;
KW inflammatory disorder; autoimmune disorder; arthritis; diabetes.
XX
XX Homo sapiens.
OS
XX WO200164955-A1.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 01-MAR-2001; 2001WO-0506572.
PF
XX
XX 02-MAR-2000; 2000US-0517467.
PR
XX
XX (ISIS-) ISIS PHARM INC.
PA
XX
XX Popoff I, Cowseert LM;
PI
```

```
XX WPI; 2001-602570/68.
DR N-PSDB; AAS45586.
XX
PT Antisense compound useful for treating hyperproliferative,
PT neurological, inflammatory and autoimmune disorders and diabetes
PT inhibits human PARP -
XX
XX Example 13; Page 102-104; 168pp; English.
XX
CC The invention relates to antisense oligonucleotides targeted to human
CC PARP nucleic acid and inhibiting expression of human PARP. PARP
CC (Poly (ADP-ribose) polymerase plays an important role in chromatin
CC decondensation, DNA replication, DNA repair, gene expression, malignant
CC transformation, cellular differentiation and apoptosis. The antisense
CC oligonucleotide inhibitors are useful for inhibiting the expression of
CC PARP in human cells or tissues. They are also useful for treating a
CC human with a disease associated with PARP especially hyperproliferative
CC disorders (e.g. cancer), cellular injury resulting from oxidative stress,
CC neurological (e.g parkinsonism, meningitis-associated intracranial
CC complications and ischaemia) , inflammatory and autoimmune disorders (e.g
CC arthritis) and diabetes. The present sequence is a PARP protein,
CC the cDNA encoding which was used to design the antisense
CC oligonucleotides.
XX
SQ Sequence 534 AA;
Query Match 25.4%; Score 725.5; DB 22; Length 534;
Best Local Similarity 35.1%; Pred. No. 2.2e-57;
Matches 196; Conservative 99; Mismatches 187; Indels 77; Gaps 21;
QY 20 GPEKKKGROAGREED--PFRSTA-----EALKAIPAKRIIRVDPCTPLSSNPG--TOVY-- 70
Db 5 gkankdrtdkdgmpgrswaskrvsesvkalllkgk-avdpdec--takvgkahvyce 61
QY 71 --BDYNTLNTQNTNIENNKKFYIIQLLO-DSNRFFTCWNRWGRVGEVGO--SKINHFTRIE 126
Db 62 gndvydvmnqnlqfnnnkyyliqlleddaqrnfswmrwgrvkgmghslvacsgnln 121
QY 127 DAKKDFEKKFREKTKNNAERDFHVSHPGKYTLIEV-----QAEDAEQAEAVVKVDRGPVRT 182
Db 122 kakeifgkfldtknwnedrekfepvgkydmlgmdayatntqdeetkkeslksplkp 181
QY 183 VTKRVQPCSLDPATQKITNIFSKEMFKNTMALMDLDVKKMPGLKSLKQOIAFGFEALEA 242
Db 182 esq-----ldlrqvclikicnvqameemmemknttkaplgkltvaqikagysgslk 235
QY 243 LEBALKGPTDGGOSLELSHFYTVIPHNFHGSQPPPIINSPELLQAKKMDLLVLADIELA 302
Db 236 iedcira-gqhgralmeacnefytriphdfglrtptlirtqkelsekiqilealadiea 294
QY 303 QALQAVSEQEKTVVEVP--HPLDRDYQLLKCQLLQSLDGAPEYKVIQTYLEQTGS----NHR 358
Db 295 ik1-----vktelqspenpldghyrlhcalrpldhesyefkvisqyiqsthapthsd 348
QY 359 CPTLQHIWKVNOEGEDRFOAHSLGNRKLLWHGTNNVAAIILTSGLRI-----MPHSGG 414
Db 349 tmlldlfevekdgekeaf--edlhnrmllwhgstrmsnvwgilshglrliaheapitgy 406
QY 415 RVGKGIYFASSENSKAGYVIGMKGCAHHVGYMFLGEVALGREHHINTDNPSSL----- 466
Db 407 mfgkglyfadmsksanycfasi--lkntglllvalgcnelleanpkagellqgkh 464
QY 467 -----KSPPPGDSVIARGHTPEP-DPTQDTELELQDQVVPVQGPVPCPEPSSSFFSQ 519
Db 465 stkglgkmapssahftlngstvtplgpasdt-----gilnpgdy-----tlny 507
QY 520 SEYLIYQESOCRLRYLLEV 538
Db 508 neyivnpgvrmryllkv 526
```

```
RESULT 11
AA51174
ID AAY51174 standard; Protein: 570 AA.
XX
AC AAY51174;
XX
DT 31-MAR-2000 (first entry)
XX
DE Human brain PARP2 protein.
XX
KW PARP; poly(ADP-ribose) polymerase; human; murine; detection; sepsis;
KW diagnosis; therapy; necrosis; apoptosis; neurodegenerative illness;
KW ischemic tissue damage.
XX
OS Homo sapiens.
XX
PN W09964572-A2.
XX
PD 16-DEC-1999.
XX
PF 04-JUN-1999; 99WO-EP03889.
XX
PR 05-JUN-1998; 98DE-1025213.
XX
PA 01-MAR-1999; 99DE-1008837.
XX
PI (BADI ) BASF AG.
XX
PI Kock M, Hoeger T, Kroeger B, Otterbach B, Lubisch W, Lemaire H;
XX
DR WPI; 2000-087218/07.
XX
DR N-PSDB; AA244287.
XX
PT Novel genes and proteins, antibodies and binding partners useful in
PT diagnosis and therapy of energy deficiency associated disease
PT conditions -
XX
PS Claim 4; Page 52-54; 96pp; German.
XX
XX This invention describes novel human and murine poly(ADP-ribose)
XX polymerase (PARP) homologues, which are characterised by an amino acid
XX sequence with a functional NAD+-binding site and no zinc finger
XX sequence motif, of general formula CX2CX2MX22C (1). The nucleic acid
XX sequences, PARP homologues and antibodies are useful for analytic
XX detection of PARP homologues and for identifying PARP effectors or
XX binding partners, as well as for determining their effectiveness.
XX PARP-binding partners are useful for the diagnosis or therapy of a
XX disease condition, which is the result of a PARP protein, especially an
XX energy deficiency, which may comprise tissue damage from cell death
XX following necrosis or apoptosis. The disease condition may be chosen
XX from a neurodegenerative illness, or sepsis or ischemic tissue damage,
XX in particular neurotoxic disturbances, etc. This sequence represents the
XX human PARP2 protein used in the method of the invention.
XX
SQ Sequence 570 AA;

Query Match 25.4%; Score 725; DB 21; Length 570;
Best Local Similarity 35.4%; Pred. No. 2.7e-57;
Matches 197; Conservative 95; Mismatches 185; Indels 80; Gaps 21;

QY 22 EKKKGKAG-----REDFRSTAEALKAIPAEKRIIRVDPTCPLSSNPG-TQVY---- 70
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
46 eskkmpvaggkankdrtdt---kqdesvkaillkgk-avpdpec--takvgkahvcegn 99

QY 71 EDYNTLNTQNIENNKKFYIIQLQ--DSNRFCTCWRGVRGVEVGQ-SKINHTRLEDA 128
DB :||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
100 dvydvmnqtnlqfnnnkyyliqiledagrnfvwmrgvkgmgqhsivacsnglnka 159

QY 129 KDKPEKFKREKTKNNWAEORHVFVSHPGKYTLIEV---QAEDEAQEAIVKVDGRGPVTVT 184
DB ||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||: ||||:
160 keifqkfldtknmwedrekfvpqkydmigmdyatnctqdeetkkeslksplkpes 219

QY 185 KRVPQCSLDPATQKLTITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGFPALEALE 244
```

```
Db 220 q-----ldlrveliklicnvqameemmmknytkkaplgkltvaqikagygslkkie 273
QY 245 EALGPTDGGOSLEELSSHFTVTPHNFPGHSQPPINSPELLOAKKMDMLLVADIELAQA 304
Db 274 dcira-gqhgrralmeacnefytriphdfglrtpplirtqkelskxiqllealgdietalk 332
QY 305 LQAVSEQEKTVEEVP-HPLDRDYQLLKCOLQLDLSGAPEYKVIOTYLEQTGS---NHRCP 360
Db 333 l-----vkteqsphepldqhyrnlhcalrpldhesyefkvisqylqsthapthsdym 386
QY 361 TLQHIWKVNOGEEDRFQAHSLGNRLKLLWHGTNNAAVVAAILTSGLRIMPH-----SGGRV 416
Db 387 tldlfevekdekeaf--edlnrmllwhgsmnvwgilshglrriappeapitgymf 444
QY 417 GKGIFYASENSKSGAGYVIGMKGAHHVGMFLGEVALGREHHINTDNPSL----- 466
Db 445 gkgiyfadmsksksanycfasr--lknrgllilsevalgcnelleanpkaeqllgqkhat 502
QY 467 ----KSPPPGFDSDVIARGHTPE-DPTODTELELDGQQVVVPGQVPVPCPEFSSTFSQSE 521
Db 503 kglgkmapssahfvltngstvpplpasdt-----gilnpgy-----tllyne 545
QY 522 YLIYQESQCRRLRYLLEV 538
Db 546 yivynpqnqvmryllkv 562

RESULT 12
AAB11480
ID AAB11480 standard; Protein: 570 AA.
XX
AC AAB11480;
XX
DT 02-MAR-2001 (first entry)
XX
DE Human brain poly-ADP-ribose-polymerase protein.
XX
KW Human; poly ADP-ribose polymerase; PARP; neuroprotective; nootropic;
KW cerebroprotective; antiparkinsonian; nephrotropic; cardiant; vasotropic;
KW anticonvulsant; cytostatic; antibacterial; immunosuppressive; treatment;
KW antiinflammatory; antirheumatic; antiarthritic; antidiabetic; epilepsy;
KW 2H-phthalazin-1-one derivative; neurodegenerative disease; tumor;
KW neuronal damage; Alzheimer's disease; Huntington's disease; metastasis;
KW Parkinson's disease; ischemic damage; microinfarction; sepsis;
KW diabetes mellitus.
XX
OS Homo sapiens.
XX
PN DEL19921567-A1.
XX
PD 16-NOV-2000.
XX
PF 11-MAY-1999; 99DE-1021567.
XX
PR 11-MAY-1999; 99DE-1021567.
XX
PA (BADI ) BASF AG.
XX
PI Lubisch W, Sadowski J, Kock M, Hoeger T;
XX
DR WPI; 2001-032983/05.
XX
DR N-PSDB; AAC82090.
XX
PT Drugs for inhibiting PARP or especially homologous enzymes comprising
PT 4-substituted phthalazinone derivatives, useful e.g. for treating
PT neurodegenerative disease, ischemic damage, tumors or diabetes -
XX
PS Example A; Page 12-13; 14pp; German.
XX
CC This invention describes novel 4-substituted 2H-phthalazin-1-one
XX derivatives (I) which are used for the treatment or prophylaxis of
XX diseases associated with elevated poly-(ADP-ribose)-polymerase (PARP;
```



Db 5 gggkankrdtdkqd-----esvkalllkqk-avpdpcc--takvgkahvycegnndvyd 54  
Qy 75 CTLNQTNLENNNNFYIIQLLQ-DSNRFCTCNRWGRVGEVQ-SKINHFRLEDAKDP 132  
Db 55 vmlqotnlqfnankkyliqllddaqrnfsvmrwgrvgkmqghslvacsgnlkakeif 114  
Qy 133 EKKFRETKKNWAPDRDHFVSHPGKYTLIEV---QAEDAEQAEVVKVDRGPRVTRVRQV 188  
Db 115 qkkfldtknnwedrekfekvgkydmlqmdyatntqdeetkkeslkspkpesq--- 171  
Qy 189 PCSLDPATQKLIITNIFSKEMPKNTMALMDLVKKMPLGKLSKQIARFEALEALEALK 248  
Db 172 ---ldlrvgelikikcnvgameemmenkntkkaplgkltvagikagyqslkkiedcir 228  
Qy 249 GPTDGGOSLELSHFYTVIPHNFGHSOPPPINSPELLOAKKMDLLVLADIELAQALQAV 308  
Db 229 a-qghgralmacehfytriphdfglrtpplirtqkseksekiqllealgdleialkl--- 284  
Qy 309 SEQKTVVEVP-HPLDRDYQLKCOLQLDLSGAPEYKYVIOTYLEQTGS---NHRCPITLQH 364  
Db 285 ---vkteqlspehpldqhyrnlhcalrpldhesyefkvisqylgsthapthsdymtild 341  
Qy 365 IWKVNOGEEDRFQAHSLGKRLKLLWHTNMVAAIILTSGLRIMPH-----SGGRVGKG 420  
Db 342 lfevkdgekeaf--edlhnmllwhgsmnsnvwgilshglrlappeapitgymfgki 399  
Qy 421 YFASENSKSAGYVIGMKGAHHVGMFLGEVALGREHHINTDNPSL----- 466  
Db 400 yfadmsksanycfasr--lknrtglllsevalggcneileanpkagellgghkstkglg 457  
Qy 467 KSPPPGFDVSIARGHTEP-DPTQDTLELDGQVVVPGQVPVPCPEFSSSTFSQSEYLIY 525  
Db 458 kmapssahfvltngstvpplgpasdt-----gilnpdgy-----tlnyneiyiv 500  
Qy 526 QESQCLRYLLEV 538  
Db 501 npnqvrmyrllkv 513

RESULT 14  
AAY68835 standard; Protein; 637 AA.  
XX AC AAY68835;  
XX DT 16-MAY-2000 (first entry)  
XX DE The poly(ADP-ribose) polymerase NAP protein of Arabidopsis.  
XX KW NAP; poly(ADP-ribose) polymerase; PARP; poly(ADP-ribose) transferase;  
XX KW programmed cell death; apoptosis; growth rate; stress; cold; pathogen;  
XX KW pest; drought; heat; fungi; nematode; seed-shatter.  
XX OS Arabidopsis thaliana.  
XX FH Key Location/Qualifiers  
XX FT Misc-difference 1..138 /note= "these residues are specifically claimed in  
XX ET claim 18"  
XX PN WO200004173-A1.  
XX PD 27-JAN-2000.  
XX PF 12-JUL-1999; 99WO-EP04940.  
XX PR 17-JUL-1998; 98US-0118276.  
XX PA (PLBZ ) PLANT GENETIC SYSTEMS NV.  
XX PI Babiychuk E, Kushnir S, De Block M;  
XX DR WPI; 2000-182436/16.

DR N-PSDB; AAZ60617.  
XX Modulating cell death, growth and stress resistance in eukaryotes,  
PT specifically plants, used, e.g. to impart fungus or nematode resistance  
PT -  
XX Claim 18; Page 99-101; 126pp; English.  
XX The present sequence represents a NAP protein. This protein is a  
CC poly(ADP-ribose) polymerase (PARP) protein (also known as  
CC poly(ADP-ribose) transferase). The PARP enzyme is involved in programmed  
CC cell death or apoptosis, and is a nuclear enzyme. The NAP polynucleotide  
CC sequences can be used for modulation of programmed cell death in  
CC eukaryotic cells. The method is used, specifically in plants, to induce,  
CC or protect against, programmed cell death, depending on the extent to  
CC which PARP activity is reduced. Reducing expression of endogenous NAP  
CC class PARP only is also used to modulate programmed cell death, to  
CC increase growth rate and to produce plant cells that are more tolerant  
CC of stress (cold, chemical treatments, pathogens, pests, drought, heat,  
CC etc., or during transformation). Particular applications are generation  
CC of plants that are resistant to fungi or nematodes; are male or female  
CC sterile; or have better seed-shatter properties. The methods are also  
CC used to improve growth of transformed plant cells (and derived calli or  
CC complete plants).  
XX Sequence 637 AA;  
SQ

Query Match 24.1%; Score 688.5; DB 21; Length 637;  
Best Local Similarity 36.2%; Pred. No. 7.2e-54;  
Matches 174; Conservative 89; Mismatches 185; Indels 33; Gaps 16;

Qy 73 YNCITNQTNLENNNNFYIIQLLQ-DSNRFCTCNRWGRVGEVQSKIN-HFRLEDAKK 130  
Db 169 ydalingtnvrndnnnkffvldesdkktymvtrvgvkgskldpydswdrale 228  
Qy 131 DFEKKFRETKKNWAPDRDHFVSHPGKYTLIEVQAEDAEQAEVVKVDRGPRVTRVRQVP- 189  
Db 229 iftnkfndktnywsdrkefiphkksytlemdygkeendspvnd---ipsssevkpe 285  
Qy 190 -CSLDPATQKLIITNIFSKEMPKNTMALMDLVKKMPLGKLSKQIARFEALEALEALK 248  
Db 286 qskldtrvakfislcnvsmmaqhmmmeigynanklplgkiskstiskgyevlkrisevi- 344  
Qy 249 GPTDGGOSLELSHFYTVIPHNFGHSOPPP--INSPPELLQAKKMDLLVLADIELAQALQ 306  
Db 345 -drydrtrleelsgefycviphdfgfkmsqfvldtpqklkqiemvealgeielatkl 403  
Qy 307 AVSBEQKTVVEVPHPDLRDYQLKCOLQLDLSGAPEYKYVIOTYLEQT-GSNHRCPTLQ-- 363  
Db 404 svdpqlq-----ddplyhyqqlncgltpvgndseefsmvanymenthakthsgytveia 458  
Qy 364 HIWKVNOGEEDRFQAHSLGKRLKLLWHTNMVAAIILTSGLRIMPH-----SGGRVGKG 419  
Db 459 qlfrastaveadrfggfssknrmllwhgsrltnwagilsgglrlappeapvtgymfgkg 518  
Qy 420 IYFASENSKSAGYVIGMKGAHHVGMFLGEVALGREHHINTDNPSLKSPPPGFDVSIAR 479  
Db 519 vyfadmfsksanycya-ntgand-gvlllcevalgdmmellysdnadnlppgklstkgv 576  
Qy 480 GHTEPDPTQDTLELDGQVVVPGQVPV--PCPEFSSSTFSQSEYLIYQESQCLRYLLE 537  
Db 577 gktapnseadtie-dg---vvvplgkpvsc---skgmlllyneiyvynveqikmryviq 630  
Qy 538 V 538  
Db 631 V 631

RESULT 15  
AAB93513  
ID AAB93513 standard; Protein; 531 AA.  
XX

AC AAB93513;  
XX  
XX 26-JUN-2001 (first entry)  
XX Human protein sequence SEQ ID NO:12844.  
DE  
XX  
XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.  
KW  
XX Homo sapiens.  
OS  
XX EP1074617-A2.  
PN  
XX 07-FEB-2001.  
PD  
XX  
XX 28-JUL-2000; 2000EP-0116126.  
PF  
XX 29-JUL-1999; 99JP-0248036.  
PR 27-AUG-1999; 99JP-0300253.  
PR 11-JAN-2000; 2000JP-0118776.  
PR 02-MAY-2000; 2000JP-0183767.  
PR 09-JUN-2000; 2000JP-0241899.  
XX  
XX (HELI-) HELIX RES INST.  
XX  
XX Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
XX  
XX WPI; 2001-318749/34.  
DR  
XX  
XX Primer sets for synthesizing polynucleotides, particularly the 5602  
PT full-length cDNAs defined in the specification, and for the detection  
PT and/or diagnosis of the abnormality of the proteins encoded by the  
PT full-length cDNAs.  
XX  
XX Claim 8; SEQ ID 12844; 2537pp + CD ROM; English.  
PS  
XX  
XX The present invention describes primer sets for synthesizing 5602  
CC full-length cDNAs defined in the specification. Where a primer set  
CC comprises: (a) an oligo-dr primer and an oligonucleotide complementary  
CC to the complementary strand of a polynucleotide which comprises one of  
CC the 5602 nucleotide sequences defined in the specification, where the  
CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
CC of an oligonucleotide comprising a sequence complementary to the  
CC complementary strand of a polynucleotide which comprises a 5'-end  
CC sequence and an oligonucleotide comprising a sequence complementary to a  
CC polynucleotide which comprises a 3'-end sequence, where the  
CC oligonucleotide comprises at least 15 nucleotides and the combination of  
CC the 5'-end sequence/3'-end sequence is selected from those defined in  
CC the specification. The primer sets can be used in antisense therapy and  
CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
CC particularly full-length cDNAs. The primers are also useful for the  
CC detection and/or diagnosis of the abnormality of the proteins encoded by  
CC the full-length cDNAs. The primers allow obtaining of the full-length  
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and  
CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to  
CC AAH95893 represent human amino acid sequences; and AAH13629 to AAH13632  
CC represent oligonucleotides, all of which are used in the exemplification  
CC of the present invention.  
XX  
XX Sequence 531 AA;  
SQ

Query Match 23.9%; Score 682.5; DB 22; Length 531;  
Best Local Similarity 36.9%; Pred. No. 1.9e-53;  
Matches 184; Conservative 86; Mismatches 173; Indels 55; Gaps 18;  
QY 22 EKKKGROAG-----REED-----PFRSTA-----EALKAIPAEEKRIIRVDPTCPLSSNP 65  
DB 46 eskmpvaggkankdrtdedqdgmpgrskvsvskallllkgk-apvdpcc--takv 102  
QY 66 G-TQVY----EDNYCTLNQTNINNKKFYIIQLLO-DSNRFFTCNWRGVRGVEGQ-SK 118  
DB 103 gkahvycegnndvdmnqnlqfnnnkyyliqlleddadqrnfsvmwrgrvqgmqghs1 162

QY 119 INHFRLEDAKDFEKKFREKTKNNWAERDFVSHPGKYTLIEV-----OAEDEAQAVVK 174  
DB 163 vacsgnlmkakeifqkklfdtknnwedrekfekypgkydmqlmqdyatntqdeetkkee 222  
QY 175 VDRGPVTVTKRVQPCSLDPATQKLTNIFSKEMFKNTMALMDLDVKKMPLGLSKSQQIA 234  
DB 223 slksplkpesq-----ldlrvgelikilcnvgameemmmenkmkntckaplgkltvaqik 276  
QY 235 RGFEALEALEEALKGPTDGGQSLLELSSHFTYVIPHNFGHSOPPPINSPELLQAKKMDLL 294  
DB 277 agyqslkktiedcira-gqhgalmecacnefytriphdfgrltpplirtkeisekiqlle 335  
QY 295 VLADIELAQALQAVSEQEXTVEVP-HPLDRDYQLLKCOLQLLDSGAPEYKVTQTYLEQT 353  
DB 336 algdieiaikl-----vkteiqspehpldqyrnlhcalrpldhesyefkvisqlqst 389  
QY 354 GS---NHRCPYLOHIWKVNOEGEEDRFQAHSKLGNKRLWLGHTNMAVVAAILTSGLRIMP 410  
DB 390 haphsdytmllldifevekdgkeaf--edlhnrmllywhgsrmanvvgilshgrirap 447  
QY 411 H---SGRVRGKGIYFASSENSKSAGYVIGMCGAHHVGYMFLGEVALGREHHINTDNPSL 466  
DB 448 peapitgymfgkyifadmskskanycfasr--lkhtgilllisevalgqcnleapnka 505  
QY 467 KSPPPGFDSVIARGHTEP 484  
DB 506 egllqgkhstkglgkmap 523

Search completed: August 29, 2002, 07:57:16  
Job time: 293 sec

**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:02:15 ; Search time 36.19 Seconds  
(without alignments)  
577.744 Million cell updates/sec

Title: US-09-701-586b-6  
Perfect score: 2854  
Sequence: 1 MSLLFLAMAPKPKPWQTEG.....EVLIIQESQCLRLYLEVLH 540

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2811	98.5	533	PP03_HUMAN	Q9Y6F1 homo sapien
2	738	25.9	559	PP02_MOUSE	O88554 mus musculus
3	729.5	25.6	583	PP02_HUMAN	Q9UGN5 homo sapien
4	688.5	24.1	637	PP02_ARATH	Q11207 arabidopsis
5	681	23.9	1015	PP01_BOVIN	P18493 bos taurus
6	674.5	23.6	996	PP01_SARPE	Q11208 sarcophaga
7	674.5	23.6	1011	PP01_CHICK	P26446 gallus gall
8	673.5	23.6	1012	PP01_CRIGR	Q9R152 cricetus
9	670	23.5	1013	PP01_HUMAN	P09874 homo sapien
10	665.5	23.3	1013	PP01_RAT	P27008 rattus norv
11	653	22.9	998	PP01_XENLA	P31669 xenopus lae
12	653	22.9	1012	PP01_MOUSE	P11103 mus musculus
13	638	22.4	994	PP01_DROME	P35875 drosophila
14	393.5	13.8	538	Y0N4_CAEEL	Q09525 caenorhabdi
15	258	9.0	1724	PP0V_HUMAN	Q9UKK3 homo sapien
16	190.5	6.7	135	PP01_ONGMA	O08824 oncorhynch
17	126	4.4	1222	YMP3_CAEEL	Q10947 caenorhabdi
18	120	4.2	1938	MYS_AEQIR	P24733 aculeipsect
19	115.5	4.0	1395	SP41_YEAST	P38904 saccharomyc
20	114	4.0	918	HXK1_BOVIN	P27595 bos taurus
21	113.5	4.0	1164	BAG_STRAG	P27951 streptococc
22	113.5	4.0	4540	DYHC_PARTE	Q27171 paramecium
23	109.5	3.8	446	GAG_OMVVS	P16900 ovine lenti
24	108.5	3.8	382	MYB_AVIMB	P01104 avian myelo
25	107	3.7	1186	SMC_BACSU	P51834 bacillus su
26	106	3.7	724	HMMR_HUMAN	O75330 homo sapien
27	106	3.7	932	YAI1A_SCHPO	Q09897 schizosacch
28	106	3.7	1057	POL_SIVAI	Q02836 simian immu
29	105.5	3.7	2025	TTC3_HUMAN	P53804 homo sapien
30	105	3.7	882	RA50_PYRFU	P58301 pyrococcus
31	104.5	3.7	641	MYB_CHICK	P01103 gallus gall
32	104.5	3.7	1403	VG22_HSV11	Q00105 ictaluriid h
33	104	3.6	1324	CUT3_SCHPO	P41004 schizosacch

34	104	3.6	1433	1	REST_CHICK	O42184 gallus gall
35	103	3.6	716	1	RRP2_IALE2	P26123 influenza a
36	103	3.6	1940	1	MYH3_RAT	P12847 rattus norv
37	102.5	3.6	640	1	MYB_HUMAN	P10242 homo sapien
38	102.5	3.6	704	1	VPS1_YEAST	P21576 saccharomyc
39	102.5	3.6	3210	1	CENF_HUMAN	P49454 homo sapien
40	102	3.6	716	1	RRP2_IALE1	P26122 influenza a
41	101.5	3.6	584	1	LIGA_HUMAN	P41214 homo sapien
42	101.5	3.6	612	1	EXO2_BPT5	P11109 bacterioph
43	101.5	3.6	874	1	SYA_PASMO	P25393 pasteurella
44	101.5	3.6	3075	1	LMA1_HUMAN	P25391 homo sapien
45	101	3.5	782	1	L100_ADE12	P36714 human adeno

ALIGNMENTS

RESULT 1

ID	PP03_HUMAN	STANDARD;	PRT;	533 AA.
AC	Q9Y6F1: Q9UG81:			
DT	16-OCT-2001 (Rel. 40, Created)			
DT	16-OCT-2001 (Rel. 40, Last sequence update)			
DE	Poly [ADP-ribose] polymerase-3 (EC 2.4.2.30) (PARP-3) (NAD(+)) ADP-ribo			
DE	ribosyltransferase-3) (Poly[ADP-ribose] synthetase-3) (PADPRT-3)			
DE	(HPARP-3).			
GN	ADPRTL3 OR PARP3 OR ADPRT3.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal brain;			
RX	MEDLINE=99263509; PubMed=10329013;			
RA	Johansson M.;			
RT	"A human poly(ADP-ribose) polymerase gene family (ADPRTL): cDNA			
RT	cloning of two novel poly(ADP-ribose) polymerase homologues.";			
RL	Genomics 57:442-445(1999).			
RP	[2]			
RN	SEQUENCE OF 75-533 FROM N.A.			
RC	TISSUE=Kidney;			
RA	Ansoerge W., Winkner U., Gassenhuber J., Wiemann S.;			
RT	Submitted (MAY-1995) to the EMBL/GenBank/DBJ databases			
CC	-1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor -			
CC	nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.			
CC	-1- SUBCELLULAR LOCATION: Nuclear (Potential).			
CC	-1- TISSUE SPECIFICITY: Widely expressed; the highest levels are in			
CC	the kidney, skeletal muscle, liver, heart and spleen; also			
CC	detected in pancreas, lung, placenta, brain, leukocytes, colon,			
CC	small intestine, ovary, testis, prostate and thymus.			
CC	-1- SIMILARITY: BELONGS TO THE PARP FAMILY.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/			
CC	or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; AF083068; AD29855.1; -			
DR	EMBL; AL050034; CAB43246.1; -			
DR	HSSP; P26446; 1A26.			
DR	InterPro; IPR001290; PARP.			
DR	InterPro; IPR004102; PARP_reg.			
DR	Pfam; PF00644; PARP; 1.			
DR	Pfam; PF02877; PARP_reg; 1.			
DR	Transferase; Glycosyltransferase; NAD; Nuclear protein;			
KW	ADP-ribosylation.			
FT	DOMAIN 14 20 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).			
FT	CONFLICT 80 80 K -> N (IN REF. 2).			

FT CONFLICT 171 171 A -> G (IN REF. 2).  
FT CONFLICT 411 411 K -> E (IN REF. 2).  
SQ SEQUENCE 533 AA; 60117 MW; 7C0AB89E64D1B9FD CRC64;

Query Match 98.5%; Score 2811; DB 1; Length 533;  
Best Local Similarity 99.6%; Pred. No. 1.3e-188;  
Matches 531; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 8 MAPKPKPWQTEGPEKKKGQAGREEDPRSTAEALKAIPAERIIIRVDTCPCLSSNPCT 67  
DB 1 MAPKPKPWQTEGPEKKKGQAGREEDPRSTAEALKAIPAERIIIRVDTCPCLSSNPCT 60  
QY 68 QYVEDYNCNLQNIENNKKFYIIQLQDSNRFCTWNRWGRVGEVGSQKINHFRLEED 127  
DB 61 QYVEDYNCNLQNIENNKKFYIIQLQDSNRFCTWNRWGRVGEVGSQKINHFRLEED 120  
QY 128 AKKDFEKKFREKTNNWAERDFVSHPGKYTLIEVQAEDEAQAQVYKVRGPRVTYTKRV 187  
DB 121 AKKDFEKKFREKTNNWAERDFVSHPGKYTLIEVQAEDEAQAQVYKVRGPRVTYTKRV 180  
QY 188 QPCSLDPATOKLTNIFSKEMFKNTWALMDLVKKPLGKLSKQQTARGFEALEALEAL 247  
DB 181 QPCSLDPATOKLTNIFSKEMFKNTWALMDLVKKPLGKLSKQQTARGFEALEALEAL 240  
QY 248 KGPTDGGQSLSEELSSHFTYVIPHNFHGSOPPPINSPPELLQAQKMDLLVLADIQAALQA 307  
DB 241 KGPTDGGQSLSEELSSHFTYVIPHNFHGSOPPPINSPPELLQAQKMDLLVLADIQAALQA 300  
QY 308 VSGQEKTEVEVPHPLDRDYQLLQDLSGAPYKVQTYLEQTSNHRCPYTLQHIWK 367  
DB 301 VSGQEKTEVEVPHPLDRDYQLLQDLSGAPYKVQTYLEQTSNHRCPYTLQHIWK 360  
QY 368 VNOGEGEDRFQAHKSLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGIFYASENS 427  
DB 361 VNOGEGEDRFQAHKSLGNRKLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGIFYASENS 420  
QY 428 KSAGYVIGMKCGAHGYMFLGEVAGLGRHHINTDPSLKSPPGPPGDSVIARGHTPEPDPT 487  
DB 421 KSAGYVIGMKCGAHGYMFLGEVAGLGRHHINTDPSLKSPPGPPGDSVIARGHTPEPDPT 480  
QY 488 QDTELELDGQVVVPGQVPVPCPEFSSSTFSQSEYLIYQESQRLYLLVHL 540  
DB 481 QDTELELDGQVVVPGQVPVPCPEFSSSTFSQSEYLIYQESQRLYLLVHL 533

RESULT 2  
PPO2\_MOUSE STANDARD; PRT; 559 AA.  
AC 088554; Q99N29;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Poly [ADP-ribose] polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-ribo-  
syltransferase-2) (Poly[ADP-ribose] synthetase-2) (ADPRT-2)  
DE (mpARP-2).  
GN ADPRTL2 OR PARP2 OR ADPRT2.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Embryo;  
RX MEDLINE=99252755; PubMed=10364231;  
RA Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiou F., Decker P.,  
RA Muller S., Hoyer T., Menissier-de Murcia J., de Murcia G.M.;  
RT "PARP-2, A novel mammalian DNA damage-dependent poly(ADP-ribose)  
polymerase.";  
RL J. Biol. Chem. 274:17860-17868(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=129SV;

RX MEDLINE=21179160; PubMed=11133988;  
RA Ame J.-C., Schreiber V., Fraulob V., Dolle P., de Murcia G.M.,  
RA Niedergang C.P.;  
RT "A bidirectional promoter connects the poly(ADP-ribose) polymerase 2  
(PARP-2) gene to the gene for RNase P RNA.";  
RL J. Biol. Chem. 276:11092-11099(2001).  
RN [3]  
RP SEQUENCE OF 9-559 FROM N.A.  
RC STRAIN=129/SV X C57BL/6;  
RX MEDLINE=99268466; PubMed=10338144;  
RA Berghammer H., Ebner M., Marksteiner R., Auer B.;  
RT "ADPRT-2: a novel mammalian polymerizing(ADP-ribose)transferase gene  
related to truncated PADPRT homologues in plants and Caenorhabditis  
elegans.";  
RL FEBS Lett. 449:259-263(1999).  
CC - FUNCTION: HAS DNA-DEPENDENT POLY(ADP-RIBOSE) POLYMERASE ACTIVITY.  
CC - SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE.  
CC - CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribo-yl}(N)-acceptor -  
nicotinamide + {ADP-D-ribo-yl}(N+1)-acceptor.  
CC - SUBCELLULAR LOCATION: Nuclear.  
CC - TISSUE SPECIFICITY: Widely expressed; the highest levels were in  
testis followed by ovary.  
CC - INDUCTION: By high levels of DNA-damaging agents.  
CC - SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC  
CC EMBL: AJ007780; CAA07679.1; -.  
CC EMBL: AF191547; AAK13253.1; -.  
CC EMBL: AF072521; AAC25415.1; ALT\_INIT.  
CC HSSP: P26446; IA26.  
CC MGD: MGI:1341112; Adprt2.  
CC InterPro: IPR001290; PARP.  
CC Pfam: PF00644; PARP; 1.  
CC Pfam: PF02877; PARP\_reg; 1.  
CC Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation.  
FT DNA\_BIND 1 65 POTENTIAL.  
FT DOMAIN 66 559 NAD-BINDING (BY SIMILARITY).  
FT DOMAIN 3 9 NUCLEAR LOCALIZATION SIGNAL 1ST PART  
(POTENTIAL).  
FT DOMAIN 33 39 NUCLEAR LOCALIZATION SIGNAL 2ND PART  
(POTENTIAL).  
FT CONFLICT 82 82 L -> V (IN REF. 2).  
FT CONFLICT 177 177 V -> I (IN REF. 2).  
FT CONFLICT 486 486 R -> Q (IN REF. 2).  
SQ SEQUENCE 559 AA; 63396 MW; E0AEDAE412C1445 CRC64;

Query Match 25.9%; Score 738; DB 1; Length 559;  
Best Local Similarity 34.3%; Pred. No. 4.4e-44;  
Matches 187; Conservative 103; Mismatches 191; Indels 64; Gaps 19;

QY 23 KKGROA-GREEDPFRSTAEALKAIPAERIIIRVDTCPCLSSNPCTQVY----EDYNCTL 77  
DB 42 QKGPVAGGKADRTKDNRSVTKLLKQK-APVDECAAKLQK-AHVYCEGDVDVDMVL 99  
QY 78 NOTNINNKKFYIIQLLO-DSNRFCTWNRWGRVGEVQ-SKINHFRLEDAKDFEKK 135  
DB 100 NOTNLFNKKFYIIQLLEDDAQRNFSVNRWGRVCKTGOHSLVTCGDLNKAKEIFQKK 159  
QY 136 FREKTKNNWAERDFVSHPGKYTLIEV---QAEDEA---QEAQVYKVRGPRVTYTKRVQ 188  
DB 160 FLDTKNNWEDRENFEKVPCKYDMLQMDYAASTQDSKTEKEETLQPE----- 207  
QY 189 PCSLDPATOKLTNIFSKEMFKNTWALMDLVKKPLGKLSKQQTARGFEALEALEAL 248

```
Db 208 -SOLDLRVQELLKLCINVTMEEMEMTKYDKRAPLGKLTVAQIKAGYQSLKKTEDCIR 266
Qy 249 GPTDGGOSLEELSSHYFTVPHNFGHSOPPPINSPELLQAKKMDLLVLADIOLAQAOAV 308
Db 267 A-QGHRALVEACNEFTYRIPDFGLSPVPVIRTEKSLDKVKKLEALGDIET--ALKLV 323
Qy 309 SEQKTEVEVPPLDRDYQLLQKQQLDLSGAPEYKVIQTYLEQT--GSNHR--CPTLQHI 365
Db 324 KSERQGLE---HPLDQHYRNLCALRPLDHSNEFKVISQYLQSTHAPTHKDYTWTLDDV 380
Qy 366 WKVNBQGEEDRQAKSKLGNKRLKLLHGTNMAVVAIILISGLIMPH-----SGRVRGKGIY 421
Db 381 FEVEKEGEKAPR--EDLPNRLMLHGSRLSNWVGLSHGLRVAPPEAPITGYMFGKGIY 438
Qy 422 FASENSKAGYVIGMKGAHHGVYMFGLGEVALGREHHINTDNPSLKSPPPGDSVIARGH 481
Db 439 FADMSKSNANYCFASR--LKNLTGLLLSEVALQCNELLEANPKAQLLGRGHKSTKGMOK 496
Qy 482 TEPDPTQDTELELDGQVVVPOGPVPCPEFSS-----TFQSEYLIYQESQCLRL 533
Db 497 MAPSPA-----HEFTLNGSTVPLGPASDGTGLNPEGTYLNAVNEFIVSPNQVRM 546
Qy 534 YLLEV 538
Db 547 YLLKI 551

RESULT 3
PPO2_HUMAN
ID PPO2_HUMAN STANDARD; PRT; 583 AA.
AC Q9UGN5; Q9Y6C8; Q9NUV2; Q9UMR4;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Poly (ADP-ribose) polymerase-2 (EC 2.4.2.30) (PARP-2) (NAD(+) ADP-
DE ribosyltransferase-2) (Poly[ADP-ribose] synthetase-2) (pADPRT-2)
DE (hPARP-2).
GN ADPRTL2 OR PARP2 OR ADPRT2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 2).
RC TISSUE=Fetal brain;
RX MEDLINE=99292755; PubMed=10364231;
RA Ame J.-C., Rolli V., Schreiber V., Niedergang C., Apiau F., Decker P.,
RA Muller S., Hoeger T., Menissier-de Murcia J., de Murcia G.M.;
RT "PARP-2, a novel mammalian DNA damage-dependent poly(ADP-ribose)
RT polymerase."
RL J. Biol. Chem. 274:17860-17868(1999).
RN [2]
RP SEQUENCE OF 2-583 FROM N.A. (ISOFORM 1).
RC TISSUE=Fetal brain;
RX MEDLINE=99263509; PubMed=10329013;
RA Johansson M.;
RT "A human poly(ADP-ribose) polymerase gene family (ADPRTL): cdna
RT cloning of two novel poly(ADP-ribose) polymerase homologues."
RL Genomics 57:442-445(1999).
RN [3]
RP SEQUENCE OF 22-583 FROM N.A. (ISOFORM 1).
RC TISSUE=Fibroblast; PubMed=10338144;
RX MEDLINE=99268466; PubMed=10338144;
RA Bernhammer H., Eder M., Marksteiner R., Auer B.;
RT "pADPRT-2: a novel mammalian polymerizing(ADP-ribose)transferase gene
RT related to truncated pADPRT homologues in plants and Caenorhabditis
RT elegans."
RL FEBS Lett. 449:259-263(1999).
RN [4]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Placenta;
RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,
```

```
RA Nishikawa T., Nagai K., Sugano S., Ishibashi T., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoka S., Ishii S., Kawai Y.,
RA Saito K., Yamamoto J., Wakamatsu A., Nakamura Y., Nagehara K.,
RA Masuho Y., Kanehori K.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: HAS DNA-DEPENDENT POLY(ADP-RIBOSE) POLYMERASE ACTIVITY.
CC SEEMS TO PLAY A ROLE IN THE RESPONSE TO DNA DAMAGE (By
CC similarity).
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor -
CC nicotinamide + [ADP-D-riboseyl](N+1)-acceptor.
CC -!- SUBCELLULAR LOCATION: NUCLEAR (By similarity).
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; may be
CC produced by alternative splicing.
CC -!- TISSUE SPECIFICITY: WIDELY EXPRESSED. THE HIGHEST LEVELS ARE IN
CC THE BRAIN, HEART, PANCREAS, SKELETAL MUSCLE AND TESTIS; ALSO
CC DETECTED IN KIDNEY, LIVER, LUNG, PLACENTA, OVARY AND SPLEEN;
CC LEVELS ARE LOW IN LEUKOCYTES, COLON, SMALL INTESTINE, PROSTATE AND
CC THYMUS.
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AJ236912; CAB65088.1; -
DR EMBL; AF085734; AAD29857.1; ALT_INIT.
DR EMBL; AJ236876; CAB41505.2; ALT_INIT.
DR EMBL; AK001980; BAA92017.1; ALT_TERM.
DR HSP; P26446; I426
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation; Alternative splicing.
FT DOMAIN 1 88 POTENTIAL.
FT DNA_BIND 86 583 NAD-BINDING (BY SIMILARITY).
FT DOMAIN 4 7 NUCLEAR LOCALIZATION SIGNAL 1ST PART
FT DOMAIN 35 40 NUCLEAR LOCALIZATION SIGNAL 2ND PART
FT DOMAIN 68 80 MISSING (IN ISOFORM 2).
FT VARSPLIC 447 447 P -> H (IN REF. 2).
FT CONFLICT 481 481 N -> H (IN REF. 4).
SQ SEQUENCE 583 AA; 66205 MW; 5B7AE8AE531836AF CRC64;

Query Match 25.6%; Score 729.5; DB 1; Length 583;
Best Local Similarity 35.4%; Pred. No. 1.8e-43;
Matches 201; Conservative 95; Mismatches 184; Indels 87; Gaps 22;

Qy 22 EKKKGKQAG-----REED-----PFRSTA-----EAKKAIPAERKRIIVDPTCLSSNP 65
Db 46 EKKKMPVAGGKANKORTEDKQGMGRSWASRVSSESVKALLKGR-APVDPEC--TAKV 102
Qy 66 G-TQVY----EDYNCTLNTQNTENNNNKFIITLLQ-DSNREFTCNWRNMGVGEVQ-SK 118
Db 103 GKHYVCEGNDYDVNLQNTNLFQNNKYYLLQLEDDAQRNFSVMRGRVKGQHS 162
Qy 119 INHFRLEDAKKDFKKFKREKTNNAERHVFVSHPGKYTLIEV-----QAEDAQEAQVVK 174
Db 163 VACSGNLNKAKEIFQKKFLDKTKNNWEDREKEPKVPGKYDMLQMDYATNTQDEETKKEE 222
Qy 175 VDRGPVRIIVTKRVQPCSLDPATQKLTITNIFSKEMPKNTMALMDLDVKMKPLGKLSKQITA 234
Db 223 SLKSPKPSQ-----LDLRVQELLKLCINVTMEEMEMTKYDKRAPLGKLTVAQIK 276
Qy 235 RGFEALEALEALKGPTDGGQSLSEELSSHFTVIPHNFCHSQPPPTNSPELLQAKKMDL 294
```

Db 277 AGYQSLKKIEDCIRA-GOGRALMEACNEFYTRIPHDGFLRTPPLIRTKQELSEKIOLLE 335  
QY 295 VLADIELAQAALQVSEQETVEVP-HPLDRDYQLLCKQLDLSGAPYKVIQTYLEQT 353  
Db 336 ALGDIETIAIKL-----VKTELQSPHLDQHYRNLCALRPDLHESYFVISOYLQST 389  
QY 354 GS---NHRCPPTLQHTWKVNOGEEDRFQASHKGLNKRLLWHTNMVAAITLTSGLRMP 410  
Db 390 HAPTHSDYTWLLDLFEVKEKGEAEFR--EDLNRMLLWHSRMSVWVGLSLGLRIAP 447  
QY 411 H-----SGRVRGKGYIFASENSKSGAGYVIGMKGAAHVGMYMFLGELVAGRGREHINTDNP 466  
Db 448 PEAPITGYMGKGYIFADMSKSNANYCFASR--LKNTGLLLLLSEVALGQCNELEAPKA 505  
QY 467 -----KSPPPGSDSVIARHTEP-DPTQDTLELGGQVQVVPQGPVCPPE 511  
Db 506 EGLQGHKSTKGLGKMAPSSAHFTVLTNGSTVPLGPASDT-----GILNPDGY----- 552  
QY 512 FSSSTFSQSEYLIYQESQCLRYLLE 538  
Db 553 ---TLNNEYIVYNQVRMYLLKV 575

## RESULT 4

PPOL\_ARATH STANDARD; PRT; 637 AA.  
AC Q11207;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PAP) (ADPRT) (NAD(+)) ADP-  
ribosyltransferase (Poly[ADP-ribose] synthetase).  
GN APP.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=95269779; PubMed=7750552;  
RA Lepinlec L., Babilchuk E., Kushair S., van Montagu M., Inze D.;  
RT "Characterization of an Arabidopsis thaliana cDNA homologue to animal  
poly(ADP-ribose) polymerase.";  
RL FEBS Lett. 364:103-108(1995).  
CC -1- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY[ADP-RIBOSYLATION]. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -1- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribose](N)-acceptor =  
CC nicotinamide + [ADP-D-ribose](N+1)-acceptor.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- SIMILARITY: BELONGS TO THE PAP FAMILY.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; Z48243; CAA88288.1; -  
CC HSP; P26446; IA26.  
CC InterPro; IPR001290; PAP.  
CC InterPro; IPR004102; PAP\_reg.  
CC Pfam; PF00644; PAP; 1.  
CC Pfam; PF02877; PAP\_reg; 1.  
CC Pfam; PF02037; PAP; 2.

DR SMART; SM00513; SAP; 2.  
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation.  
FT DNA\_BIND 1 140 POTENTIAL.  
FT DOMAIN 140 637 NAD-BINDING (BY SIMILARITY).  
FT DOMAIN 41 62 NUCLEAR LOCALIZATION SIGNAL (POTENTIAL).  
SQ SEQUENCE 637 AA; 72175 MW; 527A8F464605D127 CRC64;  
  
Query Match 24.1%; Score 688.5; DB 1; Length 637;  
Best Local Similarity 36.2%; Pred. No. 1.5e-40;  
Matches 174; Conservative 89; Mismatches 185; Indels 33; Gaps 16;  
  
QY 73 YNCFLNTNTENNKKFYIIQLQ-DSNRRFTCNRRGVRGCEVGOSKIN-HFTREDAK 130  
Db 169 YDALNTNTVNDNNKFFVLQVLESDSKKYVWYTRNGRVGVKQSKLDGPDYSDWRAIE 228  
QY 131 DFEKFKREKTKNNWAERDHFVSHFGKYTLIEVQAEQAQAVVAVKVRGVTVTKRVQP- 189  
Db 229 IFTNKFNDKTKNYSRKEFIPHPKSYTWLEMDYKKEENDSPVND--IPSSSEVKPE 285  
QY 190 -CSLDPATQKLTITNIFSKEMFKNTMALMDLVKKMPLGKLSKQOIARGFEALEEALK 248  
Db 286 QSKLDTRVAKFISLICNVSMQAQHMMEIGNANKPLPGKISKSTISKGIEVKKRISEVI- 344  
QY 249 GPTDGGQSLSELSHFYTVIPHNFHGSQPPP--INSPQLQAKMDMLLVADIQAALQ 306  
Db 345 -DRYDTRLEELSGEFTYVPHDFGFKMSQFVIDTPQKLKQKTEMVEALGEIELATKLL 403  
QY 307 AVSEQKTVVEVPHPLDRDYQLLCKQLDLSGAPYKVIQTYLEQT-GSNHRCPPTQ-- 363  
Db 404 SVDPLQ-----DDPLYYHYQQLNCGLTPVGNDSSEFSWVANYMENTHAKTHSGYTVIEA 458  
QY 364 HIWKNQGEDEDRFQASHKGLNKRLLWHTNMVAAITLTSGLRMPH----SGRVCKG 419  
Db 459 OLFRASRAVADRPFQFSSSKNRMLLWHSRLTNWGLSQGLRIAPAEAVTYMGFKG 518  
QY 420 IYFASNSKSGAGYVIGMKGAAHVGMYMFLGELVAGRGREHINTDNPSPPPGSDSVIAR 479  
Db 519 VYFADFESKSANYCYA-NTGAND-GVLLCEVALGDMNELLSDYNADNLPFGKLSKGV 576  
QY 480 GHTPEPTQDTLELGGQVQVVPQGPV--PCPEFSSSTFSQSEYLIYQESQCLRYLLE 537  
Db 577 GKYPANPSEAQTLE-DG--VVVPLGKPVRSK---SKGMLLYNEYIVYNVEQIKMYVIQ 630  
QY 538 V 538  
Db 631 V 631

## RESULT 5

PPOL\_BOVIN STANDARD; PRT; 1015 AA.  
AC P18493; Q9TS00;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PAP-1) (ADPRT) (NAD(+))  
ADP-ribosyltransferase-1 (Poly[ADP-ribose] synthetase-1).  
GN ADPRT.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
[1]  
SEQUENCE FROM N.A.  
RX MEDLINE=90382673; PubMed=2119324;  
RA Saito I., Hatakeyama K., Kido T., Ohkubo H., Nakanishi S., Ueda K.;  
RT "Cloning of a full-length cDNA encoding bovine thymus  
poly(ADP-ribose) synthetase: evolutionarily conserved segments and  
RT their potential functions.";  
RL Gene 90:249-254(1990).

[2]  
 RN SEQUENCE OF 647-714 AND 838-903 FROM N.A.  
 RX MEDLINE=88151954; PubMed=2450019;  
 RA Taniguchi T., Yamauchi K., Yamamoto T., Toyoshima K., Harada N.,  
 Tanaka H., Takahashi S., Yamamoto H., Fujimoto S.:  
 "Depression in gene expression for poly(ADP-ribose) synthetase during  
 the interferon-gamma-induced activation process of murine macrophage  
 tumor cells.\*"  
 RL Eur. J. Biochem. 171:571-575(1988).  
 CC -!- FUNCTION: POLY(ADP-RIBOSE) POLYMERASE MODIFIES VARIOUS NUCLEAR  
 CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
 CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
 CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
 CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
 CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
 CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribosyl}(N)-acceptor -  
 CC nicotinamide + {ADP-D-ribosyl}(N+1)-acceptor.  
 CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
 CC -!- SUBUNIT: HOMODIMER (Potential).  
 CC -!- SUBCELLULAR LOCATION: Nuclear.  
 CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
 CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
 CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
 CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
 CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
 CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
 CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D90073; BAA14114.1; -  
 CC EMBL: X05986; CAA30046.1; -  
 CC EMBL: X05987; CAA30047.1; -  
 CC PIR: JS0428; JS0428.  
 CC PIR: S00328; S00328.  
 CC HSP: P26446; IA26.  
 CC InterPro: IPR001357; BRCT.  
 CC InterPro: IPR001290; PARP.  
 CC InterPro: IPR004102; PARP\_reg.  
 CC InterPro: IPR001510; ZnF-PARP.  
 CC Pfam: PF00553; BRCT; 1.  
 CC Pfam: PF02877; PARP\_reg; 1.  
 CC Pfam: PF02877; PARP; 1.  
 CC Pfam: PF00645; zf-PARP; 2.  
 CC ProDom: PD004675; ZnF-PARP; 2.  
 CC SMART: SM00292; BRCT; 1.  
 CC PROSITE: PS00172; BRCT; 1.  
 CC PROSITE: PS00347; PARP\_ZN\_FINGER\_1; 2.  
 CC PROSITE: PS0064; PARP\_ZN\_FINGER\_2; 2.  
 CC Transferase: Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
 CC ADP-ribosylation; Zinc-finger; Zinc.  
 FT INIT\_MET 0 0 BY SIMILARITY.  
 FT DNA\_BIND 1 374 AUTOMODIFICATION DOMAIN.  
 FT DOMAIN 375 525  
 FT DOMAIN 386 462 BRCT.  
 FT DOMAIN 526 1015 NAD-BINDING.  
 FT ZN\_FING 20 55 PARP-TYPE.  
 FT ZN\_FING 127 164 PARP-TYPE.  
 FT DOMAIN 209 211 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
 FT DOMAIN 223 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
 FT MOD\_RES 408 408 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 414 414 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 436 436 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 446 446 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 449 449 ADP-RIBOSYL[N] (POTENTIAL).  
 FT MOD\_RES 457 457 ADP-RIBOSYL[N] (POTENTIAL).

FT	MOD_RES	472	472	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	485	485	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	489	489	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	492	492	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	514	514	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	515	515	ADP-RIBOSYL[N] (POTENTIAL).
FT	MOD_RES	521	521	ADP-RIBOSYL[N] (POTENTIAL).
SQ	SEQUENCE	1015	AA; 113355	MW; OA5FE9D9F04F5B04 CRC64;

Query Match 23.9%; Score 681; DB 1; Length 1015;  
 Best Local Similarity 33.9%; Pred. No. 9.4e-40;  
 Matches 188; Conservative 99; Mismatches 205; Indels 62; Gaps 24;

QY	7	AMAPKP	-----PWQTEGPEKKKGROAGREDDPFRSTAEALKAIPAERIRIIRVPTCP	LS 62
DB	493	AVGPKGSGAAPS	KSKGPKVEEG--TNKSEKMKLT	LKGAA-----VDPDSGLE 541
QY	63	SNPGTQYVED	-----YNCTLNQTNENNKNFYIIQLQDSNRF-FTCWNWRGRVGEV	-GQ 116
DB	542	HN--AHVLEKGGK	VFSATLGLVDIVKGTNSYKQLLEDDEKESRYWIFRSRGVGTIGS	599
QY	117	SKINHFTRL	EADAKDFEKKFKETKNNNAERDHFVSHPGKYTLIEVO-AEDEAQAQVVKV	175
DB	600	NKLEQMP	SKEDAIEHEFMKLYEETGNAHNSK-NFTKHPKKFYPLEIDYGOE--EAVKKL	656
QY	176	DRGPRVTRV	QPCSLDPATQKLTITNIFSKEMFKNTMALMDLVKKMPLGKLSKQIAR	235
DB	657	TVNP-GTKSKLP	KP-----VQNLKIMFVDSMKAMVEYIDLQKMLGKLSKROIQA	709
QY	236	GFEALEALEAL	-KGPTDGGOSLELSHFYTVIPHNFGHSOPPPINSPELLQAKKDWLL	294
DB	710	AYSLSEVQ	ALSOGSSD--SHIDLSNRFYTLIPHFGMKKPPLLNNANSVQAKVEMLD	767
QY	295	VLADIELAQA	LQAVSEQKTVVEVPHPLDRDYQLLQCOLLDLSGAPKYQVLYLQST-	353
DB	768	NLDIEVAYS	LLRGSSDDSSKD---PIDVNYEKLTDIKVVDKDSAEAEIRKYVKNTH	823
QY	354	GSNHRCTPLQ	--HIWKVNOEGEEDRFQAHSKLGNKRLHWHCTNMAVVAAILTSLGRIMPH	411
DB	824	ATTNAYDLE	VVDIFKIEREGESQRYKPKOLHNRLLWHGSRRTNFAGILUSGLRIAPP	883
QY	412	----	SGGRVKGIGYFASSENSKAGYVIGMCKGAHH---VGYMFLGVALGHEHINTDNP	464
DB	884	EAPVTG	YMGKGIYFADWVSKSANY-----CHTSQGDPIGLILLGEALGNWYELKHAR-	937
QY	465	SLKSPPPG	FDSDVIARHTEPDPTQDTELELDGQVVVQGVQGVPCPFSSSTFSQSEYLI	524
DB	938	HISKLPK	GKHSVKGIGKTTDPDS--ASITVDG--VEVPLGTGI-SSGVNDRCLLYNEVIV	992
QY	525	QYESOCRL	RYLLEV 538	
DB	993	YDIAQVHL	KLYLLKL 1006	

RESULT 6

PPOL\_SARPE

ID	PF0L_SARPE	STANDARD;	PRT;	996	AA.
AC	Q11208;				
DT	01-OCT-1996	(Rel. 34, Created)			
DT	01-OCT-1996	(Rel. 34, Last sequence update)			
DT	16-OCT-2001	(Rel. 40, Last annotation update)			
DE	Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)	ADP-			
DE	ribosyltransferase) (Poly[ADP-ribose] synthetase).				
OS	Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).				
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OX	NCBI_TaxID=7386;				
RP	[1]				
RN	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.				
RX	MEDLINE=94170813; PubMed=8125121;				
RA	Masutani M., Nozaki T., Hitomi Y., Ikejima M., Nagasaki K.,				





PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE (BY similarity).

-! CATALYTIC ACTIVITY: NAD(+) + {ADP-D-riboseyl}(N)-acceptor = nicotinamide + {ADP-D-riboseyl}(N+1)-acceptor.

-! COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN (BY SIMILARITY).

-! SUBUNIT: HOMODIMER (Potential).

-! SUBCELLULAR LOCATION: Nuclear.

-! MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN AVERAGE CHAIN LENGTH OF 20-30 UNITS.

-! SIMILARITY: BELONGS TO THE PARP FAMILY.

-! SIMILARITY: CONTAINS 1 BRCT DOMAIN.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

EMBL; AF168781; AAD45917.1; -.  
HSSP; P26446; 1A26.  
InterPro; IPR001357; BRCT.  
InterPro; IPR004102; PARP\_reg.  
InterPro; IPR001510; Znf-PARP.  
Pfam; PF00533; BRCT; 1.  
Pfam; PF02877; PARP\_reg; 1.  
Pfam; PF00645; zf-PARP; 2.  
ProDom; PD004675; Znf-PARP; 2.  
SMART; SM00292; BRCT; 1.  
PROSITE; PS0172; BRCT; 1.  
PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
PROSITE; PS50084; PARP\_ZN\_FINGER\_2; 2.  
Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
Adp-riboseylation; Zinc-finger; Zinc.  
INIT\_MET 0 0 BY SIMILARITY.  
FT DNAS\_BIND 1 371 BY SIMILARITY.  
FT DOMAIN 372 522 AUTOMODIFICATION DOMAIN (BY SIMILARITY).  
FT DOMAIN 386 475 BRCT.  
FT DOMAIN 523 1012 NAD-BINDING (BY SIMILARITY).  
FT ZN\_FING 20 55 PARP-TYPE.  
FT ZN\_FING 124 161 PARP-TYPE.  
FT DOMAIN 206 208 NUCLEAR LOCALIZATION SIGNAL 1ST PART (POTENTIAL).  
FT DOMAIN 220 225 NUCLEAR LOCALIZATION SIGNAL 2ND PART (POTENTIAL).  
FT MOD\_RES 406 406 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 412 412 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 434 434 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 436 436 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 443 443 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 444 444 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 455 455 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 470 470 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 483 483 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 487 487 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 490 490 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 511 511 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 512 512 ADP-RIBOSYL[N] (POTENTIAL).  
FT MOD\_RES 518 518 ADP-RIBOSYL[N] (POTENTIAL).  
FT SEQUENCE 1012 AA; 112400 MW; 8D617C4DBF0CB07 CRC64;

Query Match 23.6%; Score 673.5; DB 1; Length 1012;  
Best Local Similarity 33.8%; Pred. No. 3.1e-39;





```
DR EMBL; M29781; AAA51663.1; JOINED.
DR EMBL; M29783; AAA51663.1; JOINED.
DR EMBL; M29784; AAA51663.1; JOINED.
DR EMBL; M29785; AAA51663.1; JOINED.
DR EMBL; M29544; AAA51663.1; JOINED.
DR EMBL; M29782; AAA51663.1; JOINED.
DR EMBL; M56141; CAA39606.1; -.
DR EMBL; M56141; CAA39606.1; JOINED.
DR EMBL; M60436; AAA60000.1; -.
DR PIR; A28901; A26901.
DR PIR; A28498; A28498.
DR PIR; A29725; A29725.
DR PIR; A35635; A35635.
DR PIR; A33321; A33321.
DR PIR; B33321; B33321.
DR PIR; A39976; A39976.
DR PIR; S14010; S14010.
DR HSSP; P26446; 1A26.
DR Aarhus/Ghent-2DPAGE; 1620; NEPHGE.
DR MIT; 173870; -.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.

Query Match      23.5%; Score 670; DB 1; Length 1013;
Best Local Similarity 33.6%; Pred. No. 5.5e-39;
Matches 186; Conservative 101; Mismatches 195; Indels 72; Gaps 23;

QY 7  AMAPKPKPWOTEGPEKKGR-----QAGREEDP---FRSTALKAIPAERKIIRVDPT 58
Db 501 ALSKSKGQVKEGINKSEKMKLTGKGAADVDPDGLSEAHVLE----- 546

QY 59 CPLSSNPQTQVYEDYNTLQNTNIENNKNFYIQLQD---SNRFTCWNRWGRVGEV-G 115
Db 547 -----KGGKV---FSATLGLVDIVKGTNSYKYLQLEDDEKENRYW-IFRSWGRVGTG 596

QY 116 QSKINHTRLEDAKQPEKFRKTKNNAERDHFVSHPCGYTLIEVQ-AEDQAQAVVK 174
Db 597 SNLQEMPSKEDAEHFMKUYEERTGNHWSK-NFTYKPKFYPLEIDYQDE--EAVKK 653

QY 175 VDRGPVTVTKRVPQCSLDPATQKLTINIFSKEMFKNTMALMDLVKMKPLGLSKSQOIA 234
Db 654 LTVPN-CTKSKLKP-----VQDLKMIQFVESMKKAWVEYEDLQKMKPLGLSKRQIQ 706

QY 235 RGFEEALEEALKGPTDQGSLEELSHFYTVIPNFGHSQPPPIINSPELLQAKKDMLL 294
Db 707 AAYSILSEVQAVSQGSSDSQIL-DLSNRFYTLPHDFGMMKPLPLNADSVQAKVEMLD 765

QY 295 VLADIELAQALQAVSEQEKVVEVPHPDLDRDYOLLKQLQLDLSGAPYKVIQTYLEOT- 353
Db 766 NLLDIEVAYSLRGSSDDSKD-----PIDVNYEKLKTDIKVDVDRDSEAEILIKYVKNTH 821

QY 354 GSNHRCTPLQ--HIWKVQGEEDRFQAHKGLGNKRLKLLHGTNNMVAAILTSGLRIMPH 411
Db 822 ATTHNAYDLEVIDFKTEREGECQRYKPFQKLNRLHGHGSRRTNFGILSQGLRIAPP 881

QY 412 -----SGRVRKGIYFASENSKSGAGYVIGMKCGAHH-----VGYMFLGEVALGREHINTDNP 464
Db 882 EAPVTGYMFGKGIYFADMVSKSANY-----CHTSQGDPIGLILGVALGNMVELKHAS- 935

QY 465 SLKSPPPGFSVLTARGHTEPQDTQDTELELDGQGVVVPQGPVPCPEFSSSTFSQSYLI 524
Db 936 HSKLPKGSVSGKLGKTTTDPDS--ANISLDG--VDVPLGTGI-SSGVNTSLLYEYIV 990

QY 525 YQESQCRLYLLEV 538
Db 991 YDIAQVNLKYLKL 1004

RESULT 10
```

Db	821	HATTHNAYDLEVIDIFKIERGESQRYKPPRQLHNRLLMHGSRRTTFAGILSUGLRIAP	880
QY	411	H-----SGRRYKGIYFASENSKAGYVIGMKCGAAH-----VGYMFLGEVALGREHHINTDN	463
Db	881	PEAPVTGYMGKGIYFADMYVSKSANY-----CHTSQGDPIGLLGEVALGNMYELKHAS	935
QY	464	PSLKSPPPGDSVIARTGHTPEDPDTQDTELELDGQGVVVGQGPVPCPEFSSTFSQSEYL	523
Db	936	-HISKLPKGHSHVKGKLGKTKAPDPS--ASITLDG--VEVPLGTGIP-SGVNDTCILLYNEYI	989
QY	524	IYQESQCRRLRYLLEV	538
Db	990	VYDIAQVNLYKLLKL	1004

RESULT 11

PPOL\_XENLA

ID

PPOL\_XENLA

STANDARD;

PRT;

998 AA.

AC

P31669;

DT

01-JUL-1993 (Rel. 26, Created)

DT

01-JUL-1993 (Rel. 26, Last annotation update)

DT

16-OCT-2001 (Rel. 40, Last annotation update)

DE

Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)

DE

ribosyltransferase) [Poly(ADP-ribose) synthetase] (Fragment).

OS

Xenopus laevis (African clawed frog).

OC

Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;

OC

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;

OC

Xenopodinae; Xenopus.

OX

NCBI\_TaxID=8355;

RN

[1]

RN

SEQUENCE FROM N.A.

RP

TISSUE=Ovary;

RA

Saulier-Le Drian B.M.;

RL

Thesis (1992), University of Rennes, France.

RN

[2]

RP

SEQUENCE OF 742-876 FROM N.A.

RX

MEDLINE=93277538; PubMed=8503897;

RA

Ozawa Y., Uchida K., Uchida M., Ami Y., Kushida S., Okada N.,

RA

Miwa M.;

RT

"Isolation of cDNAs encoding the catalytic domain of poly(ADP-ribose)

RT

polymerase from Xenopus laevis and cherry salmon using heterologous

RT

oligonucleotide consensus sequences.";

RL

Biochem. Biophys. Res. Commun. 193:119-125(1993).

CC

-!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR

CC

PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT

CC

ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT

CC

CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND

CC

TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR

CC

EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.

CC

-!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-ribosyl](N)-acceptor =

CC

nicotinamide + [ADP-D-ribosyl](N+1)-acceptor.

CC

-!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.

CC

-!- SUBCELLULAR LOCATION: Nuclear.

CC

-!- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN OVARY, OOCYTES,

CC

AND BRAIN. LOW IN LIVER.

CC

-!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO

CC

AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND

CC

FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF

CC

THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN

CC

AVERAGE CHAIN LENGTH OF 20-30 UNITS.

CC

-!- SIMILARITY: BELONGS TO THE PARP FAMILY.

CC

-!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.

CC

-----

CC

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC

between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC

the European Bioinformatics Institute. There are no restrictions on its

CC

use by non-profit institutions as long as its content is in no way

CC

modified and this statement is not removed. Usage by and for commercial

CC

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC

or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

-----

DR

EMBL; Z12139; CAA78126.1; -

DR

EMBL; D13810; BAA02966.1; -

```
DR PIR; S31735; S31735.
DR HSP; P26446; 1A26.
DR InterPro; IPR001357; BRCT.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00533; BRCT; 1.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; zf-PARP; 2.
DR ProDom; PD004675; Znf-PARP; 2.
DR SMART; SM00292; BRCT; 1.
DR PROSITE; PS0172; BRCT; 1.
DR PROSITE; PS00347; PARP_ZN_FINGER_1; 2.
DR PROSITE; PS00864; PARP_ZN_FINGER_2; 2.
KW Transferase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;
KW ADP-ribosylation; Zinc-finger; Zinc.
FT NON_TER 1 1
FT DNA_BIND <1 356 AUTOMODIFICATION DOMAIN.
FT DOMAIN 357 507 BRCT.
FT DOMAIN 369 445 NAD-BINDING.
FT DOMAIN 508 998
FT ZN_FING 8 43 PARP-TYPE.
FT ZN_FING 111 148 PARP-TYPE.
FT DOMAIN 193 195 NUCLEAR LOCALIZATION SIGNAL 1ST PART.
FT DOMAIN 207 212 NUCLEAR LOCALIZATION SIGNAL 2ND PART.
FT MOD_RES 391 391 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 397 397 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 419 419 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 428 428 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 429 429 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 445 445 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 447 447 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 454 454 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 467 467 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 471 471 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 477 477 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 495 495 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 496 496 ADP-RIBOSYL[N] (POTENTIAL).
FT MOD_RES 503 503 ADP-RIBOSYL[N] (POTENTIAL).
FT CONFLICT 746 746 Q -> E (IN REF. 2).
SQ SEQUENCE 998 AA; 111126 MW; F5A25E4A3366BAE7 CRC64;

Query Match 22.9%; Score 653; DB 1; Length 998;
Best Local Similarity 32.5%; Pred. No. 8.2e-38;
Matches 177; Conservative 102; Mismatches 214; Indels 52; Gaps 20;

QY 10 PKPKPWQTEGPEKKKGAGREEDPPRSTAEALKAIPAEKRIIRVDPTCLSSNPQTQV 69
DB 480 PSSGPVAGKSSGKVKKEGSKNKKMLVKGAAIDPDSSEL---EDSCHVLETGG--- 533
QY 70 YEDYNTLNQNTENNKKFYIQLQLO-DSNRRFTCNWRGRVGEV-GQSKINHFTRLD 127
DB 534 -KIFSATLGLVDITRGNTSYKQLQIEHDSRYSWFRSGRVGTGIGSKKLEEMSSKD 592
QY 128 AKKDFEKKFKKNNNAERDHFVSHPGKYTLTLEQVDEAQAQVVKVDRGVPRTVTKRV 187
DB 593 AIEHFLNYQDKTGNAW-HSPNTKYPKKFYPLEIDYQGE-EDVVKLSVG-AGTKSKLA 649
QY 188 QPCSLDPATQKLITNIPFSKEMFNKMTALMDLDVKKMPLGKLSQQIARGFALEALEAL 247
DB 650 KP-----VQELIKLIPDVESMKKAMWFEIDLQKMPGLKLSKRQIQSAYSILSQVQAV 703
QY 248 KGPTDGGQSLSE-----LSHFYTVTPHNFGHSQPPINSPPELLQAKKMLMLADIELA 302
DB 704 -----SESLSEARLLDSNQFTYTLIPDFGCMKKPPLNNLEYIQARKVQMLDNLDEVA 757
QY 303 QAL--QAVSQEKTVEVPHPLDRDYOLLKCOLQLDLSGAPEYKVIQTYLEQT-GSNHRC 359
DB 758 YSLRGAGDGEK-----DPIDVKYEKIKTDIKVAKDSESRIDYVKNTHADTHNA 811
QY 360 PTLQ--HIWKVNOEGEDRFQAHSKLGNRKLWLHGTNMAYVAAITLTSGLRIMPH----SG 413
```

```
DB 812 YDLEVLFIKIDREGEYQRYKPKFKLHNRQLLWHGSRRTTNFAGILSQGLRIAPDVTG 871
QY 414 GRVKGIIYFASSENSKAGYVIGMKGAHVGYMFLGEVALGRHHINTDNPISLKSPPGF 473
DB 872 YMFEGKGIYFADMYSKSANYCHAMP--GSPICGLILLGVALGNMHELKAAISOITKL-PKG 928
QY 474 DSVIARTGTEPDPTQDTELELDGQVVVPGQVPVPCPEFSSTFSQSEYLIYQESQCLR 533
DB 929 HSVKGLGRTAPDFS--ATVQLDG--VDVPLGKGTSA-NISDTSLLYNEYIYVDIAQVNLK 983
QY 534 YLLEV 538
DB 984 YLLKL 988

RESULT 12
PPOL_MOUSE
ID PPOL_MOUSE STANDARD; PRT; 1012 AA.
AC P11103; Q9JLX4; Q9QVQ3;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Poly [ADP-ribose] polymerase-1 (EC 2.4.2.30) (PARP-1) (ADPRT) (NAD(+))
DE ADP-ribosyltransferase-1 (Poly[ADP-ribose] synthetase-1) (msPARP).
GN ADPRT OR ADPRT1 OR ADPRP.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID:10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC STRAIN=BXSB;
RX MEDLINE=89263780; PubMed=2498841;
RA Huppi K., Bhatia K., Siwarski D., Kliman D., Cherney B., Smulson M.;
RT "Sequence and organization of the mouse poly (ADP-ribose) polymerase
RT gene.";
RL Nucleic Acids Res. 17:3387-3401(1989).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM 2), AND CHARACTERIZATION.
RC STRAIN=129/Sv X C57BL/6; TISSUE=Fibroblast;
RX MEDLINE=20270268; PubMed=10809783;
RA Sallmann F.R., Vodenicharov M.D., Wang Z.-Q., Poirier G.G.;
RT "Characterization of sPARP-1. An alternative product of PARP-1 gene
RT with poly(ADP-ribose) polymerase activity independent of DNA strand
RT breaks.";
RL J. Biol. Chem. 275:15504-15511(2000).
RN [3]
RP KNOCK-OUT.
RX MEDLINE=96007847; PubMed=7578427;
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jaeger S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "On the biological role of the nuclear polymerizing NAD(+)-protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostellium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
CC -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.
CC -!- CATALYTIC ACTIVITY: NAD(+) + {ADP-D-ribose}(N)-acceptor =
CC nicotinamide + {ADP-D-ribose}(N+1)-acceptor.
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.
CC -!- SUBUNIT: HOMODIMER (Potential).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- ALTERNATIVE PRODUCTS: 2 isoforms; a long form (shown here) and a
CC short form/sPARP-1; may be produced by alternative initiation.
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN
```

CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -I- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -I- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X14206; CAA32421.1; --  
CC EMBL; AF126717; AAF61293.1; ALT\_INIT.  
CC PIR; S04200; S04200.  
CC HSP; P26446; 1A26.  
CC MGD; MGI:1340806; Adp1.1.  
DR InterPro; IPR001357; BRCT.  
DR InterPro; IPR001290; PARP.  
DR InterPro; IPR004102; PARP\_reg.  
DR InterPro; IPR001510; Znf-PARP.  
DR Pfam; PF00533; BRCT.1.  
DR Pfam; PF00644; PARP.1.  
DR Pfam; PF02877; PARP\_reg.1.  
DR Pfam; PF00645; zif-PARP; 2.  
DR ProDom; PD004675; Znf-PARP; 2.  
DR SMART; SM00292; BRCT.1.  
DR PROSITE; PS00172; BRCT.1.  
DR PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 2.  
DR PROSITE; PS00064; PARP\_ZN\_FINGER\_2; 2.  
DR Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
KW ADP-ribosylation; Zinc-finger; Zinc; Alternative initiation.  
FT INIT\_MET 0 0  
FT CHAIN 1 1012  
FT POLY [ADP-RIBOSE] POLYMERASE-1, LONG  
FT ISOFORM.  
FT CHAIN 521 1012  
FT POLY [ADP-RIBOSE] POLYMERASE-1, SHORT  
FT ISOFORM.  
FT FOR SHORT ISOFORM.  
FT DNAS\_BIND 521 521  
FT DOMAIN 1 371  
FT DOMAIN 372 522  
FT DOMAIN 384 460  
FT DOMAIN 523 1012  
FT ZN\_FING 20 55  
FT ZN\_FING 124 161  
FT DOMAIN 206 208  
FT DOMAIN 220 225  
FT MOD\_RES 406 406  
FT MOD\_RES 412 412  
FT MOD\_RES 434 434  
FT MOD\_RES 436 436  
FT MOD\_RES 443 443  
FT MOD\_RES 444 444  
FT MOD\_RES 447 447  
FT MOD\_RES 455 455  
FT MOD\_RES 483 483  
FT MOD\_RES 487 487  
FT MOD\_RES 490 490  
FT MOD\_RES 511 511  
FT MOD\_RES 512 512  
FT MOD\_RES 518 518  
FT MOD\_RES 590 590  
FT CONFLICT 607 607  
FT CONFLICT 611 611  
FT CONFLICT 628 628  
FT CONFLICT 678 678  
FT CONFLICT 702 702  
FT CONFLICT 716 716  
FT CONFLICT 757 757  
FT CONFLICT 856 856  
FT CONFLICT 981 981  
FT CONFLICT 1012 1012  
FT SEQUENCE 1012 AA; 112968 MW; 4354C3E5F01B9439 CRC64;

Query Match 22.9%; Score 653; DB 1; Length 1012;  
Best Local Similarity 32.3%; Pred. No. 8.4e-38;  
Matches 186; Conservative 102; Mismatches 211; Indels 76; Gaps 23;  
QY 2 SLLFLAMAPKPKW---VQTSGPE-----KKKGROAGREEDPFRSTAEALKAIP 47  
DB 467 SLQDLLSAHSLSPGAEVKAEPGEVAPRGKSAAPSKSKGCFKEEGVNS-----517  
QY 48 AEKRI-----IRVDPCTPLSSNPQTQVED---YNCITLQNTNINNKKFYIQLLQ 96  
DB 518 -EKRMKLTGGAADPDPSGLEHS--AHVLEKGGKVFATLGLVDIVKGTNSYKQLLE 574  
QY 97 DSNRP-FTCWNRWGRVGEV-GQSKINHTRLEDAKKDEKPEKTKNNWAERDHFVSHP 154  
DB 575 DDKESRYWIFRSWGLGTIVIGSNLQEQMPSKEEAQVQPMKLYEETGNAMHSK-NFTKYP 633  
QY 155 GKYTLIEVQ-AEDEAQEAQVAVKDRGPVTVTKRVQPCSLDPATOKLITNIFSKEMFKWTM 213  
DB 634 KKFYPLEIDYGDDE--EAVKKLTVKP-GTKSKLPKP-----VOELVGMIFDVDSMKKAL 684  
QY 214 ALMDLVKMKPLGKLSQOIARGFEALKEALKGPTDGGQGLEELSSHEVTVIPHNFG 273  
DB 685 VEYEDIQKMPGLKLSRRQIOAAYSILSEVQVPSQGSSESQLL-DLSNRYETLIPHDFG 743  
QY 274 HSQPPINSPPELLQAKDMLLVLADELALQAVSEQKTVVEVPHPLDRDYQLLKCOL 333  
DB 744 MKKPELLNADSVQAKVEMLDNLLDIEVAYSILLRGSSDSSKD-----PIDVNYEKLKTDI 799  
QY 334 QLLDSGAPEYKVIQTYLEQT-GSNHRCPTLQ--HIWKVQGEEDRFOAHSKLGKRLKLM 390  
DB 800 KVVDRDSEAEVIRKYKNTHTATTNAYDLEVIDIFKTEREGESQRYKPFQLHNRLLM 859  
QY 391 HGTNMAVVAALTSGLRIMPH---SGGRVKGIVFASSENSKAGYVIGMCKGAHH---V 443  
DB 860 HGRSTTFNAGLSOGLRIAPPEAPVTGYMFGKGYIFADWVSKSANY-----CHTSQGDPI 914  
QY 444 GYMFGEVALGRRHHINTNPSPKPPGSDSVIARGHTPDPTODTELELDGQGVVVPQ 503  
DB 915 GLIMLGEVALGNMYELKHAS-HISKLPKGKHSVGLGKTTDPDSASITL-----GVEVPL 969  
QY 504 GQVPCPEFFSSTSQSEYLYIQESQRLRYLLEV 538  
DB 970 GTGIP-SGVNDTALLYNEYIVYDIAQVNLKYLKL 1003  
RESULT 13  
PPOL\_DROME  
ID PPOL\_DROME STANDARD; PRT; 994 AA.  
AC P35875; Q9W5Q5; Q9W5S1;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Poly [ADP-ribose] polymerase (EC 2.4.2.30) (PARP) (ADPRT) (NAD(+)) ADP-  
DE ribosyltransferase) (Poly[ADP-ribose] synthetase).  
OS PARP OR CG17696/CG17718.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93234521; PubMed=8475096;  
RA Uchida K., Hanai S., Ishikawa K.-I., Ozawa Y.-I., Uchida M.,  
RA Sugimura T., Miwa M.;  
RT "Cloning of cDNA encoding Drosophila poly(ADP-ribose) polymerase:  
RT leucine zipper in the auto-modification domain.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:3481-3485(1993).  
RN [2]  
RP SEQUENCE FROM N.A.; ALTERNATIVE SPLICING, DEVELOPMENTAL STAGE, AND  
RP TISSUE SPECIFICITY.  
RX STRAIN=CANTON-S;  
RX MEDLINE=98234380; PubMed=9565614;

RA Hanai S., Uchida M., Kobayashi S., Miwa M., Uchida K.;  
RT "Genomic organization of Drosophila poly(ADP-ribose) polymerase and  
RT distribution of its mRNA during development.";  
RL J. Biol. Chem. 273:11881-11886(1998).  
RN [3]  
RP SEQUENCE FROM N.A. (LONG ISOFORM).  
RX STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlike C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svrtkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
CC -!- FUNCTION: POLY[ADP-RIBOSE] POLYMERASE MODIFIES VARIOUS NUCLEAR  
CC PROTEINS BY POLY(ADP-RIBOSYL)ATION. THE MODIFICATION IS DEPENDENT  
CC ON DNA AND IS INVOLVED IN THE REGULATION OF VARIOUS IMPORTANT  
CC CELLULAR PROCESSES SUCH AS DIFFERENTIATION, PROLIFERATION, AND  
CC TUMOR TRANSFORMATION AND ALSO IN THE REGULATION OF THE MOLECULAR  
CC EVENTS INVOLVED IN THE RECOVERY OF CELL FROM DNA DAMAGE.  
CC -!- CATALYTIC ACTIVITY: NAD(+) + [ADP-D-riboseyl](N)-acceptor =  
CC nicotinamide + {ADP-D-riboseyl}(N+1)-acceptor.  
CC -!- COFACTOR: ZINC, CONTAINS TWO MOLES OF ZINC PER MOLE OF PROTEIN.  
CC -!- SUBCELLULAR LOCATION: Nuclear.  
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
CC SHORT FORM. ARE PRODUCED BY ALTERNATIVE SPLICING.  
CC -!- TISSUE SPECIFICITY: EXPRESSED IN ADULT FEMALE OOCYTES, ANAL PLATES  
CC OF STAGE 12 EMBRYOS AND IN CELLS AROUND THE CENTRAL NERVOUS SYSTEM  
CC IN LATER EMBRYOS.  
CC -!- DEVELOPMENTAL STAGE: EXPRESSED BOTH MATERNALLY AND ZYGOTICALLY IN  
CC EMBRYOS, PUPAE AND ADULTS. EXPRESSION IS HIGHEST IN EMBRYOS.  
CC -!- MISCELLANEOUS: THE ADP-D-RIBOSYL GROUP OF NAD(+) IS TRANSFERRED TO  
CC AN ACCEPTOR CARBOXYL GROUP ON A HISTONE OR THE ENZYME ITSELF, AND  
CC FURTHER ADP-RIBOSYL GROUPS ARE TRANSFERRED TO THE 2'-POSITION OF  
CC THE TERMINAL ADENOSINE MOIETY, BUILDING UP A POLYMER WITH AN  
CC AVERAGE CHAIN LENGTH OF 20-30 UNITS.  
CC -!- SIMILARITY: BELONGS TO THE PARP FAMILY.  
CC -!- SIMILARITY: CONTAINS 1 BRCT DOMAIN.  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D13806; BAA02964.1; .  
CC EMBL; AF051548; AAC24518.1; .  
CC EMBL; AF051544; AAC24518.1; JOINED.  
CC EMBL; AF051545; AAC24518.1; JOINED.  
CC EMBL; AF051546; AAC24518.1; JOINED.  
CC EMBL; AF051547; AAC24518.1; JOINED.  
CC EMBL; AE002935; AAF45400.1; .  
CC EMBL; AE002666; .; NOT\_ANNOTATED\_CDS.  
CC EMBL; AE002892; AAF45445.2; ALT\_SEQ.  
CC PIR; A47474; A47474.  
CC HSSP; P26446; 1A26.  
CC FlyBase; FBgn0010247; Parp.  
CC InterPro; IPR001357; BRCT.  
CC InterPro; IPR001290; PARP.  
CC InterPro; IPR004102; PARP\_reg.  
CC InterPro; IPR001510; Znf-PARP.  
CC Pfam; PF00533; BRCT; 1.  
CC Pfam; PF00644; PARP; 1.  
CC Pfam; PF02877; PARP\_reg; 1.  
CC Pfam; PF00645; zf-PARP; 2.  
CC ProDom; PD004675; Znf-PARP; 2.  
CC SMART; SM00292; BRCT; 1.  
CC PROSITE; PS00172; BRCT; 1.  
CC PROSITE; PS00347; PARP\_ZN\_FINGER\_1; 1.  
CC PROSITE; PS00664; PARP\_ZN\_FINGER\_2; 2.  
CC Transferrase; Glycosyltransferase; NAD; DNA-binding; Nuclear protein;  
CC ADP-ribosylation; Zinc-finger; Zinc; Alternative splicing.  
CC FT DNA\_BIND 1 367  
CC FT DOMAIN 368 507 AUTOMODIFICATION DOMAIN.  
CC FT DOMAIN 380 454 BRCT.  
CC FT DOMAIN 508 994 NAD-BINDING.  
CC FT ZN\_FING 19 54 PARP-TYPE.  
CC FT ZN\_FING 123 161 PARP-TYPE.  
CC FT DOMAIN 208 210 NUCLEAR LOCALIZATION SIGNAL 1ST PART.  
CC FT DOMAIN 223 228 NUCLEAR LOCALIZATION SIGNAL 2ND PART.  
CC FT VARSPPLIC 376 564 MISSING (IN SHORT ISOFORM).  
CC SQ SEQUENCE 994 AA; 113791 MW; ACA85A270DD29E08 CRC64;  
  
CC Query Match 22.4%; Score 638; DB 1; Length 994;  
CC Best Local Similarity 32.2%; Pred. No. 9.1e-37;  
CC Matches 168; Conservative 100; Mismatches 195; Indels 58; Gaps 19;  
  
CC QY 44 KAIPAEKRI-----IRVDPTCLSSNPGTQYVED---YNTLTNQTNNIENNKKFYIQL 94  
CC DB 500 KSNPVSRTFKVKDGLAVDPDPSGLEDI--AHVYVDSNNKYSVWLGTLDIQRNKNYSYKVL 557  
CC QY 95 LQ-DSNRFTFCNWRGVRG--EVQSQSKINHFTRLEDAKKDFEKKFKTKNNAERDHFVS 152  
CC DB 558 LKADKKEKYWIFRSWGRIGTNGNIGNSKLEEDFTSEAKRNFEYADKTGNEYQORDNFVK 617  
CC QY 153 HPCKYTLIEVQAEDEAQAVKVDGVRVTTRKVPQCSLDPATOKLITNIFSKEMFKNT 212  
CC DB 618 RTGRMPTIEIQYDD--OKLVKHE-----SHFTSKLEISVQNLIKLIIFDIDSNNKT 667  
CC QY 213 MALMDLDVKKMPLGKLSKQOIARGFEALEALBKPTDGGQSLSEELSSHFTYVPIPHNF 272  
CC DB 668 LMEFHIDMDKMPGLKLSAHQIQSAVRVKYINVLGEGSNATK-LIDATNRFYTLIPHNF 726  
CC QY 273 GHSQPPPIPSPELLQAKKMLMLVLADIELAQALQANVSEQETVEEYVPHLDROYQLLKCO 332  
CC DB 727 GVQLPTLIETHQOIEDLRQMLDSLAEIYVSI-----IKSEYSDACNLDNHYAQIKTQ 782  
CC QY 333 LQQLDSGAPEYKVIQYQYLEOT--GSNHRCPFLQ--HWKVNQEGEEDRFQAHSLGNRKL 389  
CC DB 783 LVALDNKSEEFISLQYVKNTHASTHKSVDLKIYDVFKVSRQGEARFKPFKLLNRKL 842







**THIS PAGE BLANK (USPTO)**

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: August 29, 2002, 08:01:31 ; Search time 117.38 Seconds  
(without alignments)  
795.853 Million cell updates/sec

Title: US-09-701-586B-6  
Perfect score: 2854  
Sequence: 1 MSLLFLAMAPKPKPWQTEG.....EYLIQESQRLYLEVHL 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp-archaea:\*
- 2: sp-bacteria:\*
- 3: sp-fungi:\*
- 4: sp-human:\*
- 5: sp-invertebrate:\*
- 6: sp-mammal:\*
- 7: sp-mhcc:\*
- 8: sp-organelle:\*
- 9: sp-phage:\*
- 10: sp-plant:\*
- 11: sp-rodent:\*
- 12: sp-virus:\*
- 13: sp-vertebrate:\*
- 14: sp-unclassified:\*
- 15: sp-rviro:\*
- 16: sp-bacteriap:\*
- 17: sp-archeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2823	98.9	533	4	O96CG2
2	2249.5	78.8	528	11	O91IR6
3	828	29.0	612	5	O9TX06
4	694	24.3	635	10	O81294
5	670.5	23.5	653	10	O50017
6	669.5	23.5	607	13	O9PS82
7	669.5	23.5	1014	11	O921K2
8	653	22.9	607	13	O9PS81
9	636.5	22.3	945	5	O9N4H4
10	629	22.0	593	5	O9TX05
11	598.5	21.0	983	10	O9ZP54
12	598.5	21.0	1009	10	O9SJM4
13	578	20.3	969	10	O24370
14	578	20.3	980	10	O9ZSV1
15	534	18.7	727	5	O9XUA5
16	306.5	10.7	2276	5	O9TX01

17	302.5	10.6	815	10	O9SWB4
18	285.5	10.0	815	10	O9FX91
19	138.5	4.9	1327	4	O95271
20	137	4.8	1166	4	O9H2K2
21	137	4.8	1265	4	O9HAS4
22	134	4.7	1156	16	O66878
23	127	4.4	363	4	O9H8R9
24	125	4.4	1181	5	O9XZ37
25	125	4.4	1181	5	O9VBP3
26	124.5	4.4	429	11	O924M2
27	122.5	4.3	848	4	O9Y2K0
28	120.5	4.2	419	4	O9H9X9
29	120	4.2	1219	5	O9NJ23
30	120	4.2	1229	5	O9NJ22
31	120	4.2	1243	5	O9NJ21
32	120	4.2	1253	5	O9NJ20
33	120	4.2	1951	5	O17042
34	119.5	4.2	927	11	O91V71
35	119.5	4.2	946	11	O91XT8
36	119.5	4.2	1021	11	O91XU9
37	119.5	4.2	1052	11	O91XT7
38	119	4.2	2167	10	O9SSU1
39	119	4.2	2182	10	O9LW97
40	117	4.1	261	4	O9H8F2
41	116.5	4.1	1435	3	OQ3291
42	116	4.1	429	2	O9LAX7
43	115.5	4.0	523	11	O9DA63
44	115.5	4.0	744	4	O9BXV9
45	114.5	4.0	1214	10	O9FHI5

## ALIGNMENTS

RESULT 1

O96CG2 ID O96CG2 PRELIMINARY; PRT; 533 AA.  
AC O96CG2;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DE 01-DEC-2001 (Tremblrel. 19, Last annotation update)  
DE HYPOTHETICAL 60.1 KDA PROTEIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_taxid-9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-PRIMARY B-CELLS FROM TONSILS;  
RA Strausberg R.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC014260; AAH14260.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 533 AA; 60089 MW; 6296A0E439CC7767 CRC64;

Query Match 98.9%; Score 2823; DB 4; Length 533;  
Best Local Similarity 100.0%; Pred. No. 1.4e-208; Mismatches 0; Indels 0; Gaps 0;  
Matches 533; Conservative 0;  
Qy 8 MAPKPKWQTEGPEKKKGQAGREEDPFRSTAEAKAIPAEKRIIRVDPTCLSSNPCT 67  
Db 1 MAPKPKWQTEGPEKKKGQAGREEDPFRSTAEAKAIPAEKRIIRVDPTCLSSNPCT 60  
Qy 68 QVEDYNTLQNTNTENNKKFYIIQLQDSNRFTCNWRGVRGVGSKINHTRLED 127  
Db 61 QVEDYNTLQNTNTENNKKFYIIQLQDSNRFTCNWRGVRGVGSKINHTRLED 120  
Qy 128 AKKDEKFEKRETKNNWAERDFVSHPGKYTLIEVQAEDEAQEVVYVDGVRVTRKV 187  
Db 121 AKKDEKFEKRETKNNWAERDFVSHPGKYTLIEVQAEDEAQEVVYVDGVRVTRKV 180  
Qy 188 QPCSLDPATQKLITNIFSKEMPKNTMALMDLDVKKMPLGKLSKQQTARGFEALEALEAL 247

```

181 QPCSLDPATOKLTINIFSKEMFKNTMALDLDVKKMPLGKLSKQOIARGFEALESEAL 240
248 KGPTDGGOSLEELSSHYTYVIPHNGHSQPPPIPSPELLQAKDMLLVADIELAQAQA 307
241 KGPTDGGOSLEELSSHYTYVIPHNGHSQPPPIPSPELLQAKDMLLVADIELAQAQA 300
308 VSEQEKVEEVPHPDLDRDYQLLKQQLDLSGAPYKVIQTYLEQTSNHRCPFLQHIWK 367
301 VSEQEKVEEVPHPDLDRDYQLLKQQLDLSGAPYKVIQTYLEQTSNHRCPFLQHIWK 360
368 VNOEGEDRFOAHSKLGKGNRKLWHGTMNVAAILTSLGRIMPHSGRGVKGIFYASENS 427
361 VNOEGEDRFOAHSKLGKGNRKLWHGTMNVAAILTSLGRIMPHSGRGVKGIFYASENS 420
428 KSAGYVIGMKCGAHVGMFLGEVALGREHHINTDNPSLKSPPGFDSVTARGHTEPDPT 487
421 KSAGYVIGMKCGAHVGMFLGEVALGREHHINTDNPSLKSPPGFDSVTARGHTEPDPT 480
488 QDTELELDGQGVVPOGPVPCPFSSSTFSQSEYLYIQESQCRRLYLLEVLH 540
481 QDTELELDGQGVVPOGPVPCPFSSSTFSQSEYLYIQESQCRRLYLLEVLH 533

RESULT 2
Q91YR6 PRELIMINARY; PRT; 528 AA.
AC Q91YR6;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DE HYPOTHETICAL 59.4 KDA PROTEIN.
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N. A.
RA Strausberg R.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC014870; AAH14870.1;
KW Hypothetical protein.
SQ SEQUENCE 528 AA; 59413 MW; 82EFB0C498EB5F74 CRC64;
```

```

Query Match 78.8%; Score 2249.5; DB 11; Length 528;
Best Local Similarity 80.7%; Pred. No. 1.6e-164;
Matches 431; Conservative 37; Mismatches 59; Indels 7; Gaps 4;

QY 8 MAPKPKPWQTEGPEKKKGROAGREDDPFRSTAEALKAIPAEKRIIRVDPCTPLSSNPGT 67
DB 1 MAPKKASVQTEG--SKKQROGTEEDSFSTAEALRAAPADNRVIRVDPSCPSPRNPGI 58

QY 68 QVYEDYNCTLNQNIENNKKFYIIQLLQSDNSRFTWCNWRGVRGEVQSKINHTFRLD 127
DB 59 QVHEDYCTLTNQTIGNKKFYIIQLLEESREF--CWNWRGVRGEVQSKMNHFTCLD 117

QY 128 AKKDFEKKFKREKTKNNNAERDHFVSHPGKTYLIEVQAEADAQEAQVAVKVDGCPRTVTKRV 187
DB 118 AKKDFEKKFKREKTKNNKEERDRFVAQPNKTYLIEVQAEADAQEAQVAVKVDGCPRTV--V 174

QY 188 QPCSLDPATOKLTINIFSKEMFKNTMALDLDVKKMPLGKLSKQOIARGFEALESEAL 247
DB 175 KPCSLDPATONLTINIFSKEMFKNAWTLMLNDLVKKMPLGLTKQOIARGFEALESEAL 234

QY 248 KGPTDGGOSLEELSSHYTYVIPHNGHSQPPPIPSPELLQAKDMLLVADIELAQAQA 307
DB 235 KNPTDGGOSLEELSSCYTYVIPHNGHSRPPPIPSPELLQAKDMLLVADIELVQTLQA 294

QY 308 V-SEQEKVEEVPHPDLDRDYQLLKQQLDLSGAPYKVIQTYLEQTSNHRCPFLQHIW 366
DB 295 APGEERKEVEVPHPDLDRDYQLLKQQLDLSGESEYKAIQTYLKQTSNRCNPLRHVV 354
```

```

367 KVNQEGEDRFOAHSKLGKGNRKLWHGTMNVAAILTSLGRIMPHSGRGVKGIFYASEN 426
355 KVNREGGDRFOAHSKLGKGNRKLWHGTMNVAAILTSLGRIMPHSGRGVKGIFYASEN 414
427 SKAGYVIGMKCGAHVGMFLGEVALGREHHINTDNPSLKSPPGFDSVIARGHTPEDP 486
415 SKAGYVYTHHCGHQGVGMFLGEVALGREHHITIDDPKSPPGFDSVIARGQTEPDP 474
487 QDTELELDGQGVVPOGPVPCPFSSSTFSQSEYLYIQESQCRRLYLLEVLH 540
475 AQDIELELDGQGVVPOGPVPCPFSSSTFSQSEYLYIQESQCRRLYLLEIHL 528

RESULT 3
Q9TX06 PRELIMINARY; PRT; 612 AA.
AC Q9TX06;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE NAD+-PROTEIN(ADP-RIBOSYL)-TRANSFERASE (EC 2.4.2.30).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N. A.
RA MEDLINE-96007847; PubMed-7578427;
RA Auer B., Flick K., Wang Z.O., Haidacher D., Jager S., Berghammer H.,
RA Kofler B., Schweiger M., Wagner E.F.;
RT "on the biological role of the nuclear polymerizing NAD+-protein(ADP-
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and
RT inactivation of the ADPRT gene in the mouse.";
RL Biochimie 77:444-449(1995).
DR HSP: P26446; 1A26.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
SQ SEQUENCE 612 AA; 69241 MW; CB340F7A88FF2364 CRC64;
```

```

Query Match 29.0%; Score 828; DB 5; Length 612;
Best Local Similarity 38.4%; Pred. No. 3.5e-35;
Matches 218; Conservative 79; Mismatches 197; Indels 74; Gaps 19;

QY 6 LAMAPKPKPWQTEGPEKKKGROAGREDDPFRSTAEALKAIPAEKRIIRV-----DP 57
DB 82 LTRPTKTKTNTSEATK---TAASLDSSSESE-----DEKNQISVKIKGRAANDP 132

QY 58 TPLSSNPGTQVYED----YNTLNQNIENNKKFYIIQLLQ-DSNRFCTCWNWRGVRG 112
DB 133 HFPDSR---XHVYENGKDVYDATLNQTEISQNNKXYIIQLLEADDGSSYSVWNRWREG 189

QY 113 EVGQSKINHTR--LEDAKDFEKKFKREKTKNNNAERDHFVSHPGKTYLIEVQAEADAQ 170
DB 190 LKGOSSRKDFGKGLNQAIISLFCSEKYEKTKNTFTDRANFKKAVAGKYDMIELDYSTDS-- 247

QY 171 AVKVDGREG---PVRTVTKRV---QPCSLDPATOKLTINIFSKEMFKNTMALDLDVKK 222
DB 248 ---KPKNGASTATTATTTTKKVEHKKCSLDERVOELVKLIFDVKKMERTWEAKYDLKK 304

QY 223 MPLGLKLSKQOIARGFEALESEALKGPTDGGOSLEELSSHYTYVIPHNGHSQPPPIPS 282
DB 305 MPLGLKLSKQITKGYLVKQIEDVWGKGS--GESLSTLSSRFYTIIPAFGMSVPPVINT 362

QY 283 PELLOAKDMLLVADIELAQAQVSEQKTYVEVPHPLDRDYQLLKQQLDLSGAPE 342
DB 363 NOMLIERKNNLQNLADIETATNIKDSSEDS-----NILELHYAKLKTIDQILDNSCE 417

QY 343 YKVIQTYLEQTSNHRCPFLQHIWKVNOEGEDRFOAHSKLGKGNRKLWHGTMNVAAIL 402
DB 418 YKNLLLVTKVNYOGGKPTIVNIFKIDRDGEADRYKTKKHLGNRKLWHGSRILNYASII 477
```

Db	107	SNAPVSSNDDEADDNNGPPEEKKEKIVTATKGAANVDQWIPDDEIKSQYHVLRQGDV	166
Qy	73	NYCTLNOTNIENNKKFYIIQLLQ-DSNRFFTCWNRWGRVGEVQSKIN-HFTRLEDAKK	130
Db	167	YDALNOTVNRDNNKFFVLQVLESKKTIVYVYTRWGRVGVKQSKLDGPDVSDWDRATE	226
Qy	131	DFEKKFREKTKNWAARDHVPVSHPGKTYTLEVOAQEDAQEAUVVYVDRGPVTVTKRVQP-	189
Db	227	IFTNKFENDKTKNYWSORKKEPIHPKPYTWLEMDYGREENDSPVNNQ---IPSSSSEVKPE	283
Qy	190	-CSLDPATOKLITNIESKEMFKNTMALMDLDVKKMPLKLSKQOQIARGFEALEALEALK	248
Db	284	QSKLDTRVAKFISLICINVSMQAQHMIEIGVYANKPLGKISTKISYEVLKRISEVI-	342
Qy	249	GPTDGGOSLEELSHFVTVIPHNFGHSQPP--INSPELLQAKKMLVLADIELAQALQ	306
Db	343	-DRYDTRLEELSGEFTVTPHDFGFKKMSQFVIDTPQKLQKQIENVEALGEIETATKLL	401
Qy	307	AVSEQEKTVEEVPHLDROYLLKCOLLDGSAPEYKVIQTYLEOT-GSNHRCPTLQ---	363
Db	402	SVDPGQL-----DDPLVYHYQQLNCGLTPVGNDSSEFSVMYMENTHAKTHSGYTVET	456
Qy	364	HIWKVNOEGEDRFQAHKSLGNCRLKLLWHGTNMVAVALITSLGRIMPH-----SGRGVKG	419
Db	457	QLFRASRAVADRFQOFSSSKNMLLWHGSRITNWAGILISQGLRIAPPEAPVTGYMFGKG	516
Qy	420	IYFASENKSAGYVIGMKCGAHVGYMFEVALGREHINTDNPILSKSPPGFDSFVIAR	479
Db	517	VYFADMFKSANTCYA-NTGAND-GVLLICEVALGDMNELLISDYWNADNLPPEKLSKTKGV	574
Qy	480	GHTPEPDTQTELELDQQVVVPQGPV--PCPEFSSTFSFSEYLIYQESQCRLYLLE	537
Db	575	GKTAPNPSEAQTLLE-DG--VVVPLGRFVERSC---SKGMLLYNEYIYVNVBEQIKMRYVIQ	628
Qy	538	V 538	
Db	629	V 629	

```

RESULT      5
ID          050017
ID          050017 PRELIMINARY;      PRT;      653 AA.
AC          Q50017;
DT          01-JUN-1998 (Tr-EMBLrel. 06, Created)
DT          01-JUN-1998 (Tr-EMBLrel. 06, Last sequence update)
DT          01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE          POLY (ADP-RIBOSE) POLYMERASE.
GN          PARP.
OS          Zea mays (Maize).
OC          Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC          Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade
OC          Panicoidae; Andropogoneae; Zea.
OX          NCBI_TaxId=4577;
ON          [1]
RN          SEQUENCE FROM N.A.
RP          Babyichuk E., Cottrill P., Storozhenko S., Fuangthong M.,
RA          O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT          "Higher plants possess two poly(ADP-ribose) polymerases. ";
RL          Submitted (NOV-1997) to the EMBL/GenBank/DDBJ databases.
DR          EMBL; AJ222588; CAA10888.1; -.
DR          HSSP; P26446; 1A26.
DR          InterPro; IPR001290; PARP.
DR          InterPro; IPR004102; PARP_reg.
DR          InterPro; IPR003034; SAP.
DR          Pfam; PF00644; PARP; 1.
DR          Pfam; PF02877; PARP_reg; 1.
DR          Pfam; PF02037; SAP; 2.
DR          SMART; SM00513; SAP; 2.
SQ          SEQUENCE      653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;

```

DR	PR02037; SAP; 2.	23.5%;	Score 670.5;	DB 10;	Length 653;
DK	PRAM; SM00513; SAP; 2.	35.4%;	Pred. No. 4.9e-43;		
SQ	SEQUENCE 653 AA; 72995 MW; 5FD01923C4ABCD1D CRC64;				





DR HSSP; P26446; 1A26.  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR001510; Znf-PARP.  
DR Pfam: PF00644; PARP; 1.  
DR Pfam: PF00645; zf-PARP; 1.  
DR ProDom: PD004675; Znf-PARP; 1.  
DR PROSITE; PS50084; PARP\_ZN\_FINGER\_2; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 945 AA; 108006 MW; 1D0A62C954BC6AD9 CRC64;

Query Match 22.3%; Score 636.5; DB 5; Length 945;  
Best Local Similarity 34.0%; Pred. No. 3.5e-40;  
Matches 167; Conservative 84; Mismatches 189; Indels 51; Gaps 15;

QY 73 YNCTLNQTNINNNKYYIIQLQDSNR-FETCNRGRVG-EVGOSKINHFTREDAKK 130  
Db 477 YQATLSFTDHTQNNKSYIKTOLLKDDORENYVFRSGRGVTEGGNKHESYNSNEAIL 536  
QY 131 DFEKKFREKTNNWAERDHFVSPGKYTLIEVQAEQAQAVVDRGPVRTVTKRVQPC 190  
Db 537 KFQDVFHEKTKNDWIYKHKRMPGMSYVETDYSEFAQITDTEITPG-----SKILLPK 591  
QY 191 SLDPATOKLTINTIFSKEMFNTWALMDLDVKKMPGLKSKQOQIARGFEALEEAL-KG 249  
Db 592 SV-----KEVMSIFDVENKMSALKSFEMDNKMPGLRSLHQNINLAPEVLNDISDLVLK 647  
QY 250 PTGGGOSLEELSHFYTVIPNFECHSOPPPINSPELLOAKKMLLVADIELAQALQAVS 309  
Db 648 PIDASRL-DFSNNKYFIIPNFGMRVPEPIDSFHKIKERNMNLALLDKFAIDQISGG 706  
QY 310 EQEKTVEEVPHPDRDYLQLLKCOLQLDSDGAPYKVIQTYLEQT-GSNHRCPT-LQHIWK 367  
Db 707 DVPASTSLIGDPVDINQKLCIMEPLQOCCDDWNMIHOYLKNTGATHDLKVELIDILK 766  
QY 368 VNOGEEDRFOANSKGLGNKRLKHGTMNAVVAAILTSGLRIMPH-----SGRGVKGYYFA 423  
Db 767 LNRDNESSEKFRH--IGNRRLLWHGSGKMFAGILGQGLRIAPPEAPVSGFMFGKGYFA 824  
QY 424 SENSKAGYVIGMKGAHHGVYFGLGVALGR-EHHINTDNPSLKSPPGFDSVIARG-- 480  
Db 825 DMFSKSFY--CRANAKEAYLLLCUVALGNVQQLMASKNVRQTLPAGFQSVQGLGRQ 881  
QY 481 -----HTEPD-----PTQDTELDGQVVPQGVPCPEFSSSFQSEYLIYQESQ 529  
Db 882 CPREIGSYNHPDGYTIFLGLTYMQLOKQDV-----DYHLLYNEFIYVDVQ 928  
QY 530 CRLRYLLEVLH 540  
Db 929 IQLKYLVRVKM 939

RESULT 10  
Q9TX05  
ID Q9TX05 PRELIMINARY; PRT; 593 AA.  
AC Q9TX05  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE NAD+:PROTEIN(ADP-RIBOSYL)-TRANSFERASE, ADPRT.  
OS Drosophila sp. (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OC NCBI\_Taxid=7242;  
RN [1]  
RX MEDLINE=96007847; PubMed=7578427;  
RA Auer B., Flick K., Wang Z.Q., Haidacher D., Jager S., Berghammer H.,  
RA Kofler B., Schweizer M., Wagner E.F.;  
RT \*On the biological role of the nuclear polymerizing NAD+: protein(ADP-  
RT ribosyl) transferase (ADPRT): ADPRT from Dictyostelium discoideum and  
RT inactivation of the ADPRT gene in the mouse.\*;

RL Biochimie 77:444-449(1995).  
DR HSSP; P26446; 1A26.  
DR InterPro: IPR001357; BRCT.  
DR InterPro: IPR001290; PARP.  
DR InterPro: IPR004102; PARP\_reg.  
DR Pfam: PF00533; BRCT; 1.  
DR Pfam: PF00844; PARP; 1.  
DR Pfam: PF02877; PARP\_reg; 1.  
DR PROSITE; PS50172; BRCT; 1.  
SQ SEQUENCE 593 AA; 68018 MW; D9BA37E38B8E7CCD CRC64;

Query Match 22.0%; Score 629; DB 5; Length 593;  
Best Local Similarity 32.1%; Pred. No. 6.6e-40;  
Matches 167; Conservative 100; Mismatches 196; Indels 58; Gaps 19;

QY 44 KAIPAERKRI-----IRVDPCTPLSSNPQTQYVED----YNCTLNQTNINNNKYYIIQL 94  
Db 99 KMPVSRSTFKVKGDLAVDPDPSGLEDI--AHVYVDSNNKYSVVLGLTDIQRNKSYYKVQL 156  
QY 95 LQ-DSNRFETCNRGRVG-EVGOSKINHFTREDAKKDFEKKFREKTNNWAERDHFVS 152  
Db 157 LKADKKEKYWIFRSGRIGTGNIGNSKLEEDFTSESAKRNKEIYADKTGNEYEQORDNFVK 216  
QY 153 HPGKYTLIEVQAEQAQAVVDRGPVRTVTKRVQPCSLDPATOKLTINTIFSKEMFKN 212  
Db 217 RTGRMYPETQYDD--QKLVKHE-----SHFFTSKLEISVQNLKLIIDIDSMNKT 266  
QY 213 MALMDLDVKKMPGLKSKQOQIARGFEALEEALGKPTGGOSLEELSHFYTVIPNHF 272  
Db 267 LMFEFHIDMDKMPGLKLSAHQISAYRVVKEIYNVLECGSNTAK-LIDATNRFTYLIHPNF 325  
QY 273 GHSOPPEINSPPELLOAKKMLLVADIELAQALQAVSEQKTVVEEVPHPDRDYLQLLKCO 332  
Db 326 GVQLPTLIETHQOIEDLRQMLDSLAEIYVSI-----IKSEDYSDACNPLDNHYAQIKTQ 381  
QY 333 LQLDLSGAPYKVIQTYLEQT-GSNHRCPTLQ--HIWKVNOGEEDRFOANSKGLGNKRL 389  
Db 382 LVALDNSEEFSLSYVKNTHASTHKSVDLKIYDVDFKVSQGEARKEKPFKLLHNRKLL 441  
QY 390 WHGTMNAVVAAILTSLGRI---MPHSGRGVKGIVFASENSKSAGYVIGMKGAHHGVY 445  
Db 442 WHGSRITNFVGIILSHGLRIAPPEAPPTGYMGKIYFADVMVSKSANTCCTSQ--QNSTGL 499  
QY 446 MFLGEVALG-----REHIN--TDNPSLKSPPGFDSVIARGHTPDQTDTELDGQ 497  
Db 500 MLLSEVALGDMECTSAKYINKLSNNK-----HSCFGRGTMPDPTK-SVIRSDG- 548  
QY 498 QVYVPOGQVPCPEFSSSFQSEYLIYQESQRLRYLLEV 538  
Db 549 -VEIPYGETI-TDEHLKSLLYNEYIVYDVAQVNIQYLF 587

RESULT 11  
Q92P54  
ID Q92P54 PRELIMINARY; PRT; 983 AA.  
AC Q92P54  
DT 01-MAY-1999 (TrEMBLrel. 10, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OC NCBI\_Taxid=3702;  
RN [1]  
RX SEQUENCE FROM N.A.  
RA STRAIN=CV. LANDSBERG ERECTA;  
RA Doucet-chabeaud G., Kazmaier M.;  
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ131705; CAA10482.1; -.





```
RESULT 13
O24570 ID O24570 PRELIMINARY; PRT; 969 AA.
AC O24570;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DE POLY(ADP-RIBOSE) POLYMERASE.
GN PARP.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RA Babychuk E., Cottrell P., Storozhenko S., Fuangthong M.,
RA O'Farrell M., Van Montagu M., Inze D., Kushnir S.;
RT "Higher plants possess two poly(ADP-ribose) polymerases.";
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL: AJ222589; CAAL0889.1; -
DR HSSP: P26446; 1A26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR InterPro: IPR003034; SAP.
DR InterPro: IPR001510; ZnF-PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR Pfam: PF00645; zf-PARP; 2.
DR ProDom: PD004675; ZnF-PARP; 2.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS0172; BRCT; 1.
DR PROSITE: PS0064; PARP_ZN_FINGER_2; 2.
SQ SEQUENCE 969 AA; 109128 MW; EB23AC62EBC14009 CRC64;

Query Match 20.3%; Score 578; DB 10; Length 969;
Best Local Similarity 30.6%; Pred. No. 1.1e-35;
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

QY 23 KKGROAGREEDPFRSTAALKAIPAERKRIIRVDPTCLSSNPGTQVYEDYNTLNQTN 82
DB 480 KVGSRGSAVHSSGLQDTAHL-----DGKSI---YNATLNMSDL 516
QY 83 ENNNFYIITQLL-QDSNRFCTWNRGRV--EVGSKINHTFRLDADKDFEKKFREK 139
DB 517 ALGVNSYIVLQIEQDDGSECYVFRKWRGVGSEKIGQKLEEMSKTE-AIKEFKRLFLEK 575
QY 140 TKNNWAE---RDHFVSHPGKYTLIEVOAEDAQEAQAVKVDGRGVRVTRKRVQPCSLDPAT 196
DB 576 TGSWEAWECKTNFRKQPGRFYPLDVG-----YGVKAPKPKRDISEMK-SSLAPQL 625
QY 197 QKLITNIFSKEMPKNTMALMDLVKKMPLGKLSKQOIAQGFEALEALEALKGPTDGGOS 256
DB 626 LELMKLFNVETVRAAMFEINNSEMPGLKSKENIEKGFALTETQNLKDKTADQALA 685
QY 257 LEE-----LSSHFTYVTPHNFHSGSPPPINSPELLOAKKMDLLVLADIQAQAV-SE 310
DB 686 VRESLIVAASNRFFTLIP-----SIHPHITRDEDDLMKAKMKEALQDIEATSKIVGFDSD 741
QY 311 QEKTVEEVPHPLDRDYQLLKQLQLDLSGAPEYKVITQYLEQTSNHRCP-----LQ 363
DB 742 SDES-----LDDKMKMLHCDITPLAHSDEYKLEIYQL---LNTAPTHKDWISLE 790
QY 364 HIKNVQEGEDRFQAH-SKLGNRKLLHGTNNMAVVAALITSGRLTMPH-----SGGRVKG 418
DB 791 EVFSLDRDGLNRYSRKKNLHKNLMLHWSRLTNFVGLISQGLRIAPPEAPVTGYMFGK 850
QY 419 GIYFASENSKSAGYVIGMKCGA---HHVGYMFLGEVALGREHINTDNPGLKSPPPGFS 475
DB 851 GLYFADLVSKSAQY-----CYVDRNPNVGLMLLSEVALGDMYELKKAT-SMDKPPRGKHS 904
```

```
QY 476 VIARGHTEPDPTDTELELDGQOQVVVPGQVPVPCPEFSSSTFSQSEVLYIYQSCRLRYL 535
DB 905 TKGLGKTVP---LESEFVKWRDDVVVPCGKVPV-SSIRSELMYNEIVYNTSQVKMQFL 960
QY 536 LEV 538
DB 961 LKV 963

RESULT 14
Q9ZSV1 ID Q9ZSV1 PRELIMINARY; PRT; 980 AA.
AC Q9ZSV1;
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE POLY(ADP-RIBOSE) POLYMERASE (EC 2.4.2.30).
GN PARP1.
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-99026291; PubMed-9808734;
RA Mahajan P.B., Zuo Z.;
RT "Purification and cDNA cloning of maize Poly(ADP)-ribose polymerase.";
RL Plant Physiol. 118:895-905(1998).
DR EMBL: AF093627; AAC79704.1; -
DR HSSP: P26446; 1A26.
DR InterPro: IPR001357; BRCT.
DR InterPro: IPR001290; PARP.
DR InterPro: IPR004102; PARP_reg.
DR InterPro: IPR003034; SAP.
DR InterPro: IPR001510; ZnF-PARP.
DR Pfam: PF00533; BRCT; 1.
DR Pfam: PF00644; PARP; 1.
DR Pfam: PF02877; PARP_reg; 1.
DR Pfam: PF00645; zf-PARP; 2.
DR ProDom: PD004675; ZnF-PARP; 2.
DR SMART: SM00292; BRCT; 1.
DR PROSITE: PS0172; BRCT; 1.
DR PROSITE: PS0064; PARP_ZN_FINGER_2; 2.
KW Transferase; Glycosyltransferase.
SQ SEQUENCE 980 AA; 110475 MW; 9D8AED26BC37E5C1 CRC64;
```

Query Match 20.3%; Score 578; DB 10; Length 980;  
Best Local Similarity 30.6%; Pred. No. 1.1e-35;  
Matches 166; Conservative 96; Mismatches 195; Indels 86; Gaps 21;

```
QY 23 KKGROAGREEDPFRSTAALKAIPAERKRIIRVDPTCLSSNPGTQVYEDYNTLNQTN 82
DB 491 KVGSRGSAVHSSGLQDTAHL-----DGKSI---YNATLNMSDL 527
QY 83 ENNNFYIITQLL-QDSNRFCTWNRGRV--EVGSKINHTFRLDADKDFEKKFREK 139
DB 528 ALGVNSYIVLQIEQDDGSECYVFRKWRGVGSEKIGQKLEEMSKTE-AIKEFKRLFLEK 586
QY 140 TKNNWAE---RDHFVSHPGKYTLIEVOAEDAQEAQAVKVDGRGVRVTRKRVQPCSLDPAT 196
DB 587 TGSWEAWECKTNFRKQPGRFYPLDVG-----YGVKAPKPKRDISEMK-SSLAPQL 636
QY 197 QKLITNIFSKEMPKNTMALMDLVKKMPLGKLSKQOIAQGFEALEALEALKGPTDGGOS 256
DB 637 LELMKLFNVETVRAAMFEINNSEMPGLKSKENIEKGFALTETQNLKDKTADQALA 696
QY 257 LEE-----LSSHFTYVTPHNFHSGSPPPINSPELLOAKKMDLLVLADIQAQAV-SE 310
DB 697 VRESLIVAASNRFFTLIP-----SIHPHITRDEDDLMKAKMKEALQDIEATSKIVGFDSD 752
```

```

Qy 311 QKTVVEVPHLPDRDYQLLKCOLQLDLSGAPEYKVIOTYLTQTSNHRCTP-----LQ 363
Db 753 SDES-----LDDKYNKLHCDDITPLAHSDDYKLIQYL-----LNTHAPTHKWSLELE 801
Qy 364 HIWKNQGEEDRQAH-SKLGKRLKLLWHGTNMAVRAILTSGLRIMPH-----SGGRVGK 418
Db 802 EVSLDRDGEJUNKYSRYKNNJHMKLLWHGSRLLNFVGLISQGLRIAPPEAPVTGYMFGK 861
Qy 419 GIYFASSENSAGYVIGMKCGA---HVGYNMFLGEVALGREHHINTDNPSLKSPPPGDFS 475
Db 862 GLYFADLVSKSAQY-----CIVDRNNPGLMLLSEVALGDMYELKKAT-SNDKPPRGKHS 915
Qy 476 VIARHTEPDPTQTELELDGQQVVPQGPVPCPEFSSTFSSEYLIYQESQCRRLYL 535
Db 916 TKGLGKTVP---LESEFVKWRDDVVVPCGKVP--SSIRSELMTNEXIVYNTSQVKQFL 971
Qy 536 LEV 538
Db 972 LKV 974

RESULT 15
Q9XUA5
ID Q9XUA5 PRELIMINARY; PRT; 727 AA.
AC Q9XUA5;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE AC8.1 PROTEIN.
GN AC8.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McMurray A.A.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL; Z83097; CAB05448.1; -.
DR HSSP; P26446; I426.
DR InterPro; IPR001290; PARP.
DR InterPro; IPR004102; PARP_reg.
DR InterPro; IPR001510; Znf-PARP.
DR Pfam; PF00644; PARP; 1.
DR Pfam; PF02877; PARP_reg; 1.
DR Pfam; PF00645; zf-PARP; 1.
DR ProDom; PD004675; Znf-PARP; 1.
DR PROSITE; PS50064; PARP_ZN_FINGER_2; 1.
SQ SEQUENCE 727 AA; 82884 MW; 530ABA8E991FFED CRC64;

```

```

Query Match 18.7%; Score 534; DB 5; Length 727;
Best Local Similarity 28.6%; Pred. No. 1.8e-32;
Matches 162; Conservative 104; Mismatches 211; Indels 90; Gaps 20;

Qy 37 RSTAEALKAIPAEXRIIRVDPTCPPLSSNPGTQVYEDYNCTLNNTNINNKKFYIIQLQ 96
Db 182 KATGEYVEALAKGG--STEPATPASGP--TPPEAETPVLSAEGSPSSNKRPAEAETIE 238
Qy 97 -----DSNRFF--TCNWRNVRGVEYQSKI-----NHFTRELD 127
Db 239 IDGEGNPDENDFAKKRMKKEARLMEVQKKRMKQSDLLWEYRIFEREMPTDNIILRE 298
Qy 128 AKDPF-----EKKFRETKNNWAERDHFVHPGKYTLIEVAQDEAQAQVYKVRGP 179
Db 299 NEODIPEGHRTAQDFHEKTKNDWIYRKHKRMPGMFYSYVETDYS-----FVGTNNGH 353

```

```

Qy 180 VRTVT--KRVQPCS---LDPATQKLTNIFSKEMFKNTMALMDLDVKMPLKLSKQOIA 234
Db 354 KKKITPGSKITPGSKITLLPKSVKEVMSIFDVENMKSALKSFEIDVNMKPLGRLSHNQIN 413
Qy 235 RGFALEALEAL-KGPTDGGQSLEELSSHYTVIPHNFGHSQPPPINSPPELLQAKDML 293
Db 414 LAFEVLNDISDLLVKLPIDASKIL-DFSNNKYTTIIPHNFGMRVPEPIDSFHKIKEKNML 472
Qy 294 LVLADIELAQAQAVSQEKTVEEVPHPDRDYQLLKCOLQLDLSGAPEYKVIOTYLTQTS 353
Db 473 NALLDIKPAYDQISGGDVPASTSLSDPVDINRYRLKLCIMEPLQOQGGDDNMHIQYLKNT 532
Qy 354 -GSNHRCTP-LQHIWKYNQEGEEDRFQAHSKLGNKRLKLLWHGTNMAVRAILTSGLRIMPH 411
Db 533 HGATHDLKVELIDILKYNRDNESKFRH--IGNRRLWHGSGKMNFAgilGQGLRIAPP 590
Qy 412 ----SGRVKGITYPASSENSKAGYVIGMKCGAHHVGMFLGEVALGR-EHHINTDNPSL 466
Db 591 EAPVSGYMFGRGVYFADMFSSKFFY---CRANAKEEAYLLLCDVALGNVQOLMASHKNYSR 647
Qy 467 KSPPPGDFSVIARG-----HTEPD----PTQDTELELDGQQVVPVPGQPVPCPEFS 513
Db 648 QTLPAFGFSQVQGVGRQCPREIGSYKNPDGYTVPLGLTYMQLQGNV----- 694
Qy 514 SSTFSQSEYLIYQESQCRRLRYLLEVLH 540
Db 695 DYHLLYNEFIVYDQIQLKLYLVRVKM 721

```

Search completed: August 29, 2002, 08:01:34  
Job time: 366 sec

**THIS PAGE BLANK (USPTO)**

---



Db 502 ALSKSKGQVKEGINKSEKRMKLTLLKGAADVDPDSGLEHSAHVLE----- 547  
QY 59 CPLSSNPGTQVYEDYNTQNTNIENNKKFYLIQLQD--SNRFTCNWRGRVGEV-G 115  
Db 548 -----KGGKV---FSATLGLVDIVKGTNSYIKQLLEDDEKKNRYW-IFRSWGRVGTIG 597  
QY 116 QSKINHFTLEDAKDFEKKFKKTKNNNAERDHFVSHPGKYTLIEVQ-AEDEAQAQAVVK 174  
Db 598 SNKLEOMPSKEDAIEHFMKLYEKTGNAMHSK-NFTKPKFKFPLEIDYQDDE--EAVKK 654  
QY 175 VDRGPVRTVTKRVQPCSLDPATQKLTNIFSKEMFKNTMALDLDVKKMPLGLKSKQOIA 234  
Db 655 LTVNP-GTKSKLPK-----VQDLIKMIFDVESMKAMVEYEIDLQKMPGLKSKRQIQ 707  
QY 235 RGFEALEALEAKGPTDQGSLSEESSHFYVPIPHFGHSOPPPINSPELLOAKDMLL 294  
Db 708 AAYSILSEVQAVSQGSSDSQIL-DLSNRFYTLIPHDFGKMKPPLNNADSVQAKVEMLD 766  
QY 295 VLADIELAQAQAVSQEKTVEVPHPLDRDYQLLCKQLQDLSGAPEYKVIQTYLEQT- 353  
Db 767 NLLDIEVAYSLRGSSDDSKD---PIDVNYEKLKTDIKVVDRODSEEAIEIRKYVKNTH 822  
QY 354 GSNHRCPTIQ--HIWKVNOGEEDRFQAHSKLGNRKLWHGTNMVAVAILTSGLRIMPH 411  
Db 823 ATTHSAYDLEVIDIFKIEREGECQRYKPKQLHNRLLWHGSRRTTNFAGILSQGLRIAPP 881  
QY 412 ----SGRVGKGIYFASENSKSAGYVIGMCKGAHH-----VGYMPLGEVAGLGRHHINT 461  
Db 883 EAPVTGYMFGKGIYFADVMYKSANY-----YHVSQGDPIGLIILLGEVAGLGNMYELKH 934  
QY 462 DNPSLKSPPPGDSVIARGHTEPDPTQDTELELDGQOVVVPQGPVPCPEFSSTFSQSE 521  
Db 935 AS-HISRLPKGKHSVGLGKTTDPDS--ANISLDG--VDVPLGTGIGSSGVIDTFSLL-YNE 988  
QY 522 YLIYQSQOCLRYLLEV 538  
Db 989 YIYDIAQVNLKYLKLL 1005

RESULT 2  
US-08-860-886-2  
; Sequence 2, Application US/08860886  
; Patent No. 6335009  
; GENERAL INFORMATION:  
; APPLICANT: Burkle, Alexander  
; APPLICANT: Zur Hausen, Harald  
; APPLICANT: Jan-Heiner, Kupper  
; TITLE OF INVENTION: VECTORS AND VIRUSES FOR USE  
; TITLE OF INVENTION: IN GENE THERAPY  
; NUMBER OF SEQUENCES: 2  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds, LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036-2811  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: Windows  
; SOFTWARE: FastSeq for Windows Version 2.0b  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/860,886  
; FILING DATE: 03-OCT-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Coruzzi, Laura A  
; REGISTRATION NUMBER: 30,742  
; REFERENCE/DOCKET NUMBER: 8484-0028-999

; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-493-4935  
; TELEFAX: 650-493-5556  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1013 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
US-08-860-886-2

Query Match 23.5%; Score 670; DB 4; Length 1013;  
Best Local Similarity 33.6%; Pred. No. 1.7e-59;  
Matches 186; Conservative 101; Mismatches 195; Indels 72; Gaps 23;

QY 7 AMAPKPKWQTEGPEKKGR-----QAGREDP---FRSTAEALKAIPAERKRIIRDPT 58  
Db 501 ALSKSKGQVKEGINKSEKRMKLTLLKGAADVDPDSGLEHSAHVLE----- 546  
QY 59 CPLSSNPGTQVYEDYNTQNTNIENNKKFYLIQLQD--SNRFTCNWRGRVGEV-G 115  
Db 547 -----KGGKV---FSATLGLVDIVKGTNSYIKQLLEDDEKKNRYW-IFRSWGRVGTIG 596  
QY 116 QSKINHFTLEDAKDFEKKFKKTKNNNAERDHFVSHPGKYTLIEVQ-AEDEAQAQAVVK 174  
Db 597 SNKLEOMPSKEDAIEHFMKLYEKTGNAMHSK-NFTKPKFKFPLEIDYQDDE--EAVKK 653  
QY 175 VDRGPVRTVTKRVQPCSLDPATQKLTNIFSKEMFKNTMALDLDVKKMPLGLKSKQOIA 234  
Db 654 LTVNP-GTKSKLPK-----VQDLIKMIFDVESMKAMVEYEIDLQKMPGLKSKRQIQ 706  
QY 235 RGFEALEALEAKGPTDQGSLSEESSHFYVPIPHFGHSOPPPINSPELLOAKDMLL 294  
Db 707 AAYSILSEVQAVSQGSSDSQIL-DLSNRFYTLIPHDFGKMKPPLNNADSVQAKVEMLD 765  
QY 295 VLADIELAQAQAVSQEKTVEVPHPLDRDYQLLCKQLQDLSGAPEYKVIQTYLEQT- 353  
Db 766 NLLDIEVAYSLRGSSDDSKD---PIDVNYEKLKTDIKVVDRODSEEAIEIRKYVKNTH 821  
QY 354 GSNHRCPTIQ--HIWKVNOGEEDRFQAHSKLGNRKLWHGTNMVAVAILTSGLRIMPH 411  
Db 822 ATTHSAYDLEVIDIFKIEREGECQRYKPKQLHNRLLWHGSRRTTNFAGILSQGLRIAPP 881  
QY 412 ----SGRVGKGIYFASENSKSAGYVIGMCKGAHH-----VGYMPLGEVAGLGRHHINTDNP 464  
Db 882 EAPVTGYMFGKGIYFADVMYKSANY-----CHTSQGDPIGLIILLGEVAGLGNMYELKHAS- 935  
QY 465 SLKSPPPGDSVIARGHTEPDPTQDTELELDGQOVVVPQGPVPCPEFSSTFSQSEYLI 524  
Db 936 HISKLPGKHSVGLGKTTDPDS--ANISLDG--VDVPLGTGI-SSGVNDTSLLYNEYIV 990  
QY 525 YOESQCLRYLLEV 538  
Db 991 YDIAQVNLKYLKLL 1004

RESULT 3  
US-09-196-387-2  
; Sequence 2, Application US/09196387  
; Patent No. 627613  
; GENERAL INFORMATION:  
; APPLICANT: de Lange, Titia  
; APPLICANT: Smith, Susan  
; TITLE OF INVENTION: A PROTEIN THAT BINDS TO TRF1 AND METHODS  
; TITLE OF INVENTION: OF USE THEREOF  
; NUMBER OF SEQUENCES: 12  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Klauber & Jackson  
; STREET: 411 Hackensack Avenue, 4th Floor

CITY: Hackensack  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/09/196.387  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 09/095,225  
FILING DATE: June 10, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Jackson Esq., David A.  
REGISTRATION NUMBER: 26,742  
REFERENCE/DOCKET NUMBER: 600-1-230 CIP1  
TELEPHONE: 201-487-5800  
TELEFAX: 201-343-1684  
TELEX: 133521  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1327 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHETICAL: NO

US-09-196-387-2

Query Match 4.9%; Score 138.5; DB 4; Length 1327;  
Best Local Similarity 24.4%; Pred. No. 0.00012;  
Matches 73; Conservative 41; Mismatches 102; Indels 83; Gaps 16;

Qy 284 ELLQAKDMLLVADI---ELAQ-ALQAVSQEKTVEVPHPPL-----DRDYQLLKCQLQ 334  
Db 1045 DIFEQITLDVADMGHEHKEIGINAYGHRKLKIGVERLLGSGQQNTNPLYTFHCYNQ 1104  
Qy 335 ---LSDGAP---EYKVIQTLEQTSNHRCP-----TLQHWKVNQEGEDRF- 377  
Db 1105 GTILLD-APEDKEVQSVSEEQSTIREHROGNGAGIFNRVNVIRIQKVNKKLRERFC 1163  
Qy 378 -----QAHSKLGNRKLHMGTNMAVVAAILTSLG-RIMPHSGRGVKGKGYFASENSKS 429  
Db 1164 HRQKEVSEENHHNHRMLFHS--PFNATHKGFDERHAYIGMGAGIYFAENSXKS 1221  
Qy 430 AGYVIGM---KCGAH-----HVGVMFLGEVALGREHINTDNPGLKSPPGFDSVI 477  
Db 1222 NOYVYVGGGTGCPHKDRSCYICHRQLFC-RVTLGKS-FLQFTMKMAHAPGHHSVI 1279  
Qy 478 ARGHTPEPTDTELELDGQVVVPQGPVPCPEFSSTFSOSELYLYQESQCRRLYLL 536  
Db 1280 -----GRP-----SVNGLAYAEVYIRGEQAYEYLL 1306

RESULT 4

US-08-923-992A-2  
Sequence 2, Application US/08923992A  
Patent No. 6280738  
GENERAL INFORMATION:  
APPLICANT: Tai, Joseph Y.  
APPLICANT: Blake, Milan S.  
TITLE OF INVENTION: NO. 6280738-IgA Fc Binding Forms of the Group B  
TITLE OF INVENTION: Streptococcal Beta Antigens  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
STREET: 1100 New York Avenue, N.W., Suite 600

CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US/08/923.992A  
FILING DATE: 05-SEP-1997  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 60/024,707  
FILING DATE: 06-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Esmond, Robert W.  
REGISTRATION NUMBER: 32,893  
REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
TELEPHONE: (202) 371-2600  
TELEFAX: (202) 371-2540  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1164 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-923-992A-2

Query Match 4.0%; Score 113.5; DB 4; Length 1164;  
Best Local Similarity 20.2%; Pred. No. 0.034;  
Matches 107; Conservative 72; Mismatches 205; Indels 147; Gaps 25;

Qy 28 QAGREEDPFRSTAEALKAIPAFAEKRIIRDPTCTPLSSNPGCTVYEDYDYNCTLNQTNIE-NNN 86  
Db 309 QATQVKNQFLENAQKLEIQ-----PLKKEINVKLYKAMESLEQVEKELKHN 356  
Qy 87 NKFYIQLLQDSNRFFTCW-----NRWGRVGEVQSGKINHFTLEDAKDKFEKFKREK --- 139  
Db 357 SEANLEDLVAKSKETLVREYEGKLNQSKNLPKLQLEEEAHSKLQVDEFRKKFKTSEQV 416  
Qy 140 -----TKNWAERDHFVSHPGKYTLIEOVADEAQAQVAVKVDGRGPV---RTVT 184  
Db 417 TPKKRVKRDLAANENNOQKIELTVS-PENITVYE--GEDVKFTVAKSDSKTTLDFSDLL 473  
Qy 185 KRVQPCSLDPATQKLTITNIFSKEMFKNTMALMDLDVKKHPLGKLSKQQTARGFEALEALE 244  
Db 474 TKYNPSVSDRISTNYKTN-----TDNHKIAETIKNL---KLNESQTV----- 513  
Qy 245 EALKGPTGGQSLEBLSHFYTVIPHNFCHSQPPINSPELLOAKKMDMLLVADIQLAQA 304  
Db 514 -TLKAKDSGNVVEKT-----FTITVQKKEEQVP--KTPE-----QKD----- 549  
Qy 305 LQAVSEQEKTVEEVP-HPLDRDYQLLKCQLQ-LLDGAPYKVIQTYLEQTSNHRCP 362  
Db 550 -----SKTEKVPQEPKSN-----KNQLQLIISAQOELEKKAIKEL----- 589  
Qy 363 QHMKVNOEGEDRFQAHSKLGNRKLHMGTNMAVVAAILTSLGRIMPHSGRGVKGKGYF 422  
Db 590 -----MEQPEIPSNPEYGIQKSIWESQKEPIQEI-TSFKKIIGDSSSKYVTEHYF 639  
Qy 423 ASENSKAGYVIGMKGAHHVGYMFLGEVALGREHIN--TDNPSLKSPPPGFDSDVIAR- 479  
Db 640 NKYKSDFNYYQL-----HAQMEMLTRKVV-----QYMKNPDPNAEIKKI---FESDMKRT 686  
Qy 480 -----GHTPEP-----PTQDTELELDGQVVVPQGPVPCPEFS 513  
Db 687 KEDNYGSLNDALKGYFEKYFLTPFNKIKQIVDDLDKKVQEQDQAPAPIPENS 737

```

Query Match      3.9%; Score 112.5; DB 4; Length 717;
Best Local Similarity 19.3%; Pred. No. 0.019;
Matches 102; Conservative 81; Mismatches 229; Indels 117; Gaps 21;

QY      20  GPBKKGROAGREDP---FRSTAEALKATPAEKRIIRVDPCTLSSNPCTOVVEDYNC 75
      | : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db      87  GGERTRPRESQESDPEDDDVKKPALOSSVAVTSKERTRD-----LIQDN- 133

QY      76  TLNQTNIENNKKFYII-----QLLODSNRFTTCWNKRGVGEVGGQKINHTLEDAKK 130
      | : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     134  -MEKGQKRRRIIFGLMGTLOKRFQEST---VATERQKRQEIEQ-KLE--VQAEERK 186

QY     131  DFEKKFKFEKTKNNAERDHFVSHPGKVTLYIEVQAE-DEAQEAAYVKDVRGPVVTKR--- 186
      | : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     187  QVENERELFEERRAKQTELRLLEKQVELAQLOEWEHNAKIKY---IRTKTKPHLF 242

QY     187  VQCSLDPAQOKLITNFSK---EMFKNTMALMDLVKMKPLG-----KLSKQOIARGFE 238
      | : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     243  YIPGRMCPAQTKLIESSQRKNMALFGERRIEFAQINMKRPRRQSMKEKEHQVVRNEE 302

QY     239  ALEALEEALGPTDGGGQSLEELSSHFTVPIPNFGHSQPPPIINSPELLQAKDMLLVAD 298
      | : : : : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db     303  QKAEQEGKVAQRE--EELETGNHNDVIEEAGEEE-----EKETAIHSD 348

```

```

Query Match      3.98;  Score 112.5;  DB 4;  Length 1098;
Best Local Similarity 20.08;  Pred. No. 0.039;
Matches 107;  Conservative 75;  Mismatches 199;  Indels 153;  Gaps 26;

QY 28 QAGREDFPRTAALKAIPAERKIIRVDPTCPLESSNPGTVYEDYDCTLNQTNIE-NNN 86
   || : : || : || : || : || : || : || : || : || : || : || : ||
Db 272 QATVQKNOFLANAQKLKEIQ-----PLIKETNVKLYKAMSELSQVQEKELKHN 319

QY 87 NKFYIILOLDSNFFTCW----NRGRVGVGVGOSKTNHFTLEDAKDKDFEKKFEK --- 139
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 320 SEANLEDPLVAKSKETVREYEGKLNOSKNLPELKOLEEAHKSLLKOVVEDPRKKFKTSIOV 379
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Query Match 3.9%; Score 112.5; DB 4; Length 1098;

Best Local Similarity 20.0%; Pred. No. 0.039;  
Matches 107; Conservative 75; Mismatches 199; Indels 153; Gaps 26;

Qy	28	QAQREDPFRSTAEALKAIPAEKRIIRVDPCTPLSNPGTQVYEDVNCFTLNTQINIE--NNN	86
		: :    :    :    :    :    :    :    :    :    :    :    :    :    :	
Dd	272	QATQVNQNFLENAAKUKEIO-----PLIKETNVKLYKAMSESLVEQVELKH	319
Qy	87	NKFVIQTLLQDSNRFFTCW-----NRGRVGVEGQSQIINHFTLEDAKKDFEKFKFEK----	139
Dd	320	SEANLEPDAVKSETIVREYEGKUNOSKNIPELKOLEEAHSKLKVQVEDFRKFKFTEOV	379



Qy 140 -----TKNWAERDHFVSHPKYTLIEVOAEDAQAQVVKYVDRGPV---RVVT 184  
Db 380 TPKKRLKRDLAANENNOQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDLDFSL 436  
Qy 185 KRVPQCSLDPATOKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGFEALEALE 244  
Db 437 TKYNPSVSDRISTNYKTN-----TDNHKTAETIKNL---KLNESQTV----- 476  
Qy 245 EALKGPTDGGOSLELSHFYTVIPHNFGHSQPPPINSPPELLQAKKMLLVADIELAOA 304  
Db 477 -TLAKADDSGNVVEKT-----FTITVQKKEEKQVP--KTPE-----QKD----- 512  
Qy 305 LQAVSEQEKTVEEVP-HPLDRDYOLLKCOLQ-LLDGSAPEYKVIOQTYLEOTGNSHRCPTL 362  
Db 513 -----SKTEKVPQEPKSN-----KNQLELKSAAQOELEKLEKAIKEL----- 552  
Qy 363 QHIWKVNOGEDRFOAHSKLGKLNKRLKLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGIFY 422  
Db 553 -----MEQPELPSNPEYGIQKSIWESQKEPIQEI--TSFKKIIGDSSSKYVTEHYF 602  
Qy 423 ASENSKAGYVIGMKGAAHHVGYMFLGEV-ALGRE--HHIN--TDNPSLKSPPPGFDVSI 477  
Db 603 -----NKYKSHFMNYQLHAQMEMLTRKVVQYMKYDPNABEIKKI---FESDM 646  
Qy 478 AR-----GHTPEP-----PTQDTELELDGQVVVVGQVPPCPPEFS 513  
Db 647 KRTKEDNYGSLNDALKGYFEKYFLTPFNKIKQIVDDLDKVKVEDQAPAPIPENS 700

RESULT 7  
US-08-923-992A-6  
; Sequence 6, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/923,992A  
; FILING DATE: 05-SEP-1997  
; CLASSIFICATION: 536  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/024,707  
; FILING DATE: 06-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Esmond, Robert W.  
; REGISTRATION NUMBER: 32,893  
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 371-2600  
; TELEFAX: (202) 371-2540  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1128 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-923-992A-6

Query Match 3.9%; Score 110.5; DB 4; Length 1128;  
Best Local Similarity 20.0%; Pred. No. 0.066;  
Matches 106; Conservative 73; Mismatches 205; Indels 147; Gaps 25;  
Qy 28 QAGREEDPFRSTAEALKAIPAEKRIIRVDPTCPPLSSNPGTOVYEDYNCTLQNTNIE-NNN 86  
Db 273 QATQVKNQFLENAOKLKEMQ-----PLIKETNVKLYKAMSESLQOEVEKELKHN 320  
Qy 87 NKFYIIQLLODSNFFETCW-----NRWGRVGEVSGOSKINHFTRLDEAKDKDFKKFKREK--- 139  
Db 321 SEANLEDLVAKSKEITREVECKLQNSKLPKOLKEEAAHSLKQOVVEDFRKKFKTSQV 380  
Qy 140 -----TKNWAERDHFVSHPKYTLIEVOAEDAQAQVVKYVDRGPV---RVVT 184  
Db 381 TPKKRVKRDLAANENNOQKIELTVS-PENITVYE--GEDVKFTVTAKSDSKTTLDLDFSL 437  
Qy 185 KRVPQCSLDPATOKLITNIFSKEMFKNTMALMDLDVKKMPLGKLSKQOIARGFEALEALE 244  
Db 438 TKYNPSVSDRISTNYKTN-----TDNHKTAETIKNL---KLNESQTV----- 477  
Qy 245 EALKGPTDGGOSLELSHFYTVIPHNFGHSQPPPINSPPELLQAKKMLLVADIELAOA 304  
Db 478 -TLAKADDSGNVVEKT-----FTITVQKKEEKQVP--KTPE-----QKD----- 513  
Qy 305 LQAVSEQEKTVEEVP-HPLDRDYOLLKCOLQ-LLDGSAPEYKVIOQTYLEOTGNSHRCPTL 362  
Db 514 -----SKTEKVPQEPKSN-----KNQLELKSAAQOELEKLEKAIKEL----- 553  
Qy 363 QHIWKVNOGEDRFOAHSKLGKLNKRLKLLWHGTNMAVVAAILTSGLRIMPHSGRGVKGIFY 422  
Db 554 -----MEQPELPSNPEYGIQKSIWESQKEPIQEI--TSFKKIIGDSSSKYVTEHYF 603  
Qy 423 ASENSKAGYVIGMKGAAHHVGYMFLGEV-ALGRE--HHIN--TDNPSLKSPPPGFDVSIAR- 479  
Db 604 NKYKSHFMNYQL-----HAQMEMLTRKVV---QYMKYDPNABEIKKI---FESDMKRT 650  
Qy 480 -----GHTPEP-----PTQDTELELDGQVVVVGQVPPCPPEFS 513  
Db 651 KEDNYGSLNDALKGYFEKYFLTPFNKIKQIVDDLDKVKVEDQAPAPIPENS 701  
RESULT 8  
US-08-910-925-3  
; Sequence 3, Application US/08910925  
; Patent No. 6162601  
; GENERAL INFORMATION:  
; APPLICANT: Bandman, Olga  
; APPLICANT: Lal, Preeti  
; APPLICANT: Shah, Purvi  
; TITLE OF INVENTION: HUMAN PININ SPLICE VARIANT  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Incyte Pharmaceuticals, Inc.  
; STREET: 3174 Porter Drive  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/910,925  
; FILING DATE: Herewith  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:



Qy 370 QEGEDRFQAH-SKLGNRKLLW-----HGTNNVAVVAAILTSGLRIMPHSGGRV-GK 418  
Db 1267 ---KEAMRKHIDALLERKVRPSKSPRTN----AFIVESGTSIDPKTGKIRGK 1314

## RESULT 10

US-09-349-546-1  
; Sequence 1, Application US/09349546  
; Patent No. 6093569  
; GENERAL INFORMATION:  
; APPLICANT: Olszewski, N.  
; APPLICANT: Tzafir, I.  
; APPLICANT: Somers, D.A.  
; APPLICANT: Lockhart, B.  
; APPLICANT: Torbert, K.  
; TITLE OF INVENTION: Sugarcane bacilliform virus promoter  
; FILE REFERENCE: 600.369US2  
; CURRENT APPLICATION NUMBER: US/09/349,546  
; CURRENT FILING DATE: 1999-07-08  
; EARLIER APPLICATION NUMBER: US 08/694,869  
; EARLIER FILING DATE: 1996-08-09  
; EARLIER APPLICATION NUMBER: PCT/IB97/01338  
; EARLIER FILING DATE: 1997-08-13  
; NUMBER OF SEQ ID NOS: 10  
; SOFTWARE: FastSEQ for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1871  
; TYPE: PRT  
; ORGANISM: sugarcane bacilliform virus  
US-09-349-546-1

Query Match 3.8%; Score 108.5; DB 3; Length 1871;  
Best Local Similarity 21.0%; Pred. No. 0.25; Indels 177; Gaps 25;  
Matches 100; Conservative 64; Mismatches 135;

Qy 40 AEALKAIPAERRIIRVDPTCLSSNPGTQVYEDYNTLNQTNNIENNKKFYIIQLQDSN 99  
Db 919 AEAVK--PPEKK-----SNVELLAKOLLIIENSKLMEKEILIEBLN 957  
Qy 100 RFFTCWNRWGR-----VGEVQSQKINIFTRLEDAKKDFEKKFRKTKNNWA----- 145  
Db 958 KEIKAHOETKKGKELYIEEASTEVENETWKSRAELFEALYNEEVKKNKASTSSVTEGM 1017  
Qy 146 ---ERDRFVSHPGYKTLIEVQA-----EDEAQEAUV-----KVDRGPVRTYTKR 186  
Db 1018 YQVQIDHL-----RKEURVEATLEVNKVESEEEAEVMMASAVKDEMYRFPVILIEVPE 1072  
Qy 187 VQPCSLDP-----ATOKLITNIFSKEMF-----KNTMALDLDVYK----- 221  
Db 1073 VGKVLTAALDTGATSCINQVFIIEFKLQPTKFKVKGHGVNSYTKLDQVKGAKLWAG 1132  
Qy 222 ---KMPD---GKL---SKQIARGFEALEALEALKGPTDGGOSLELSHFYTVIPHN 271  
Db 1133 ENWFLPITYVGYMGMGEKTOMLGCFNMOSLA-----GGVLEGRVTTFYKYI--- 1181  
Qy 272 FGHSPQPPINSPELLOAKKMLLVADIELAQALQAVSEQE-----KTVEEVPH 320  
Db 1182 ---ASIKANEYLQAEAEELV-----ATSEQEFINRSPMSKNRLLLEEMK- 1223  
Qy 321 PLDRDY-----QLKRCQLQLDSDGAPEYKVIQTLYEQTGSNHRCPQLQHIKVN 369  
Db 1224 ---EQGYWGEDTFLAHNNKQIKCKLELN---PDLLI-----KDKPOTLLNIQK-- 1266  
Qy 370 QEGEDRFQAH-SKLGNRKLLW-----HGTNNVAVVAAILTSGLRIMPHSGGRV-GK 418  
Db 1267 ---KEAMRKHIDALLERKVRPSKSPRTN----AFIVESGTSIDPKTGKIRGK 1314

## RESULT 11

US-09-172-422-1  
; Sequence 1, Application US/09172422A  
; Patent No. 6300485

; GENERAL INFORMATION:  
; APPLICANT: Adams, Arwen E.  
; APPLICANT: Chiu, Choi Ying  
; APPLICANT: Duhl, David  
; APPLICANT: Gorman, Susan W.  
; APPLICANT: Leng, Song  
; APPLICANT: Sheffield, Val  
; APPLICANT: Welch, Juliet  
; TITLE OF INVENTION: MYOSIN IXA AND CYCLIC NUCLEOTIDE GATED  
; TITLE OF INVENTION: CHANNEL-15 (CNCC-15) POLYNUCLEOTIDES, POLYPEPTIDES,  
; FILE REFERENCE: 200130.442  
; CURRENT APPLICATION NUMBER: US/09/172,422A  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: FastSEQ for Windows Version 3.0  
; SEQ ID NO 1  
; LENGTH: 2548  
; TYPE: PRT  
; ORGANISM: Homo sapien  
US-09-172-422-1

Query Match 3.8%; Score 108.5; DB 4; Length 2548;  
Best Local Similarity 21.0%; Pred. No. 0.42;  
Matches 88; Conservative 63; Mismatches 126; Indels 143; Gaps 25;

Qy 41 EALKAIPAERKRIIRVDPTCLSSNPGTQVYEDYNTLNQTNNIENNKKFYIIQLQDSN 99  
Db 757 EILQRCKEEYSITRKNPRTPLSLDQG-----MNALNEKNQHDFTDI----- 798  
Qy 100 RFFTCWNRWGRVGEVQSQKINIFTRLEDAKKDFEKKFRKTKNNWAERDH-----FV 151  
Db 799 ---AWN---GTG---IROSRUSSGTSLLDK-----DGIFANSTSSKLLERAHGILTRKNFK 847  
Qy 152 SHPG---KYTLIEVQA-----EDEAQEAUVKVDGPRVTVTKRVQPCSLDPATQ--- 197  
Db 848 SKPALPKH-LLEVNSLKHLLTRTLQDRITKSLLLH-----KKKKPPSISAQFOASL 898  
Qy 198 -KLITNIFSKEMFNTMALDMDLVKKPLGKLSKQQTARGFEALEALB--BALKGPDTGG 254  
Db 899 SKMETLQGAEPY--FVKCIRSNAEKPLP-RFSDVLVRLQRLRYTGMLETQVIRQSGYSK 955  
Qy 255 QSLSELSHFYTVIPHNFHGSQPPPI-----SPELLOAKKMLLVADIELAQ 304  
Db 956 YSFQDFVSHFHVLLPRNI---IPSKFNIQDFRINLPDNYQVCKTMVF----- 1002  
Qy 305 LQAVSEQEK---VEEVPHP-----LDRDYQLLKCQLQLDSDGAPEYKVIQTLYEQTGS 355  
Db 1003 ---LKEQERQHLQDLHQQEVLRRILLQRFVRLLCROHFL-----HLRQAS- 1046  
Qy 356 NHRCPQLQHIWK--VNOEG-----EEDRFOAHKSLGNKRLKLWHGTNNVAVVAAILTSGLR 407  
Db 1047 ---VIIQREWRNYLNQKQVRDAAVQKDAFV-----MASAAALLQASWR 1086

## RESULT 12

US-08-923-992A-10  
; Sequence 10, Application US/08923992A  
; Patent No. 6280738  
; GENERAL INFORMATION:  
; APPLICANT: Tai, Joseph Y.  
; APPLICANT: Blake, Milan S.  
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B  
; TITLE OF INVENTION: Streptococcal Beta Antigens  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.  
; STREET: 1100 New York Avenue, N.W., Suite 600  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005

```
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-923-992A-10

Query Match 3.8%; Score 107.5; DB 4; Length 1164;
Best Local Similarity 20.4%; Pred. No. 0.14;
Matches 108; Conservative 68; Mismatches 210; Indels 143; Gaps 24;

QY 28 QAGREEDPFRSTAEALKAIPAERRIIRDVPTCLSSNPGTVQVYEDYNCTLNQTNIE--NNN 86
Db 309 QATQVKNFLNQAUKLEIQ-----PLIETNVKLYKAMSESLQVEKQKLNH 356
QY 87 NKFYIQLQDSDNRFFTCW----NRMRGVGEVQSKINHFTRLEDAKDKPEKFR---EK 139
Db 357 SQANLEDLVAKSEIVREYEGSKNQSNLPQLAQLEEAHSLKLVQVDFRFRKFTSEQV 416
QY 140 TKNWAERDFHVGHPGKYTLIEVQAEDEAQAENVKVDGPGVIVTKRVQPCSLDPATQKL 199
Db 417 TPKRKVRKD--LAANENNQQKIELTVSPEN---ITVYEG-----EDL 453
QY 200 ITNIFSEKEMFNMTMALMDLVKKMP-----LGKLSKQOIA-----RGFEALALEE 245
Db 454 KFTLTAKSDSKTLDPSDLLTKNPSVSDRISNTYKNTDNDHKAETITKNKLNESQTV 513
QY 246 ALKGPTDGGQSLSELSHFYTVIPHNFHSGQPPPIINSPELLQAKKMDLLVLADIQAAL 305
Db 514 TLKAKDSDGNVQKT-----FTITVQKKEQVP--KTPE---QKD-----549
QY 306 QAVSEQEKTVEEVP--HPLDRDYQLLKCQLQLDLSGAPEYKVIQTYLEQTSNHRCTLOH 364
Db 550 -----SKTEEKVPQEPKSDN---KNQQL-----IKSAQQQLKLEK-----584
QY 365 IKVNOBGEDEDRFOASKICLNKRLNHHGTNNVAAITLTSGLRIMPHSGRGVCKGIVFAS 424
Db 585 --AIKELMEQPEIPSPNPEYGIQISWESQEPQEI--TSFKKLIIGDSSSKYYTEHYFNK 641
QY 425 ENSKAGYVIGMCGAHHVGMFLGEVALGREHIN--TDNPLSKPPPGFDSVIAR---479
Db 642 YKSDFMNYQL-----HAQMEMLTRKVV-----QYINKYPDPAEIKKI---FESDMKRTKE 688
QY 480 ---GHTPEP-----PQDTELELDGQVVVVPQVPCPEFS 513
Db 689 DNYGSLNDALKGYFEKIFYLTPFNKIQIVDDLDKVKVEQDQAPAPIPENS 737

RESULT 13
US-09-045-360-2
; Sequence 2, Application US/09045360
; Patent No. 6207880
```

```
;
; GENERAL INFORMATION:
; APPLICANT: Jens Kossmann
; APPLICANT: Ruth Lorberth
; TITLE OF INVENTION: PLANTS WHICH SYNTHESIZE A MODIFIED STARCH,
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FISH & NEAVE
; STREET: 1251 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10020
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPA)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/045,360
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/EP96/04109
; FILING DATE: 19-SEP-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19547733.2
; FILING DATE: 20-DEC-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: DE 19534759.5
; FILING DATE: 19-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Haley Jr., James F.
; REGISTRATION NUMBER: 27,794
; REFERENCE/DOCKET NUMBER: GFB-5
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-596-9000
; TELEFAX: 212-596-9090
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1464 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-045-360-2

Query Match 3.7%; Score 105.5; DB 4; Length 1464;
Best Local Similarity 20.8%; Pred. No. 0.34;
Matches 91; Conservative 73; Mismatches 193; Indels 81; Gaps 18;

QY 44 KATPAE-KRIIRDVPTCLSSNPGTVQVYEDYNCTLN---QTNIE---NNNKFYIQLLQ 96
Db 61 KKTPEMKKRAFSSPHAVLTDTSSLEAKFSLGGNIELQVDVRRPPTSGDVSVDFQVTN 120
QY 97 DSNRFTTCNWRNVRGVGEVQSKINHFTRLEDAKDKPEKFRKTKNNWAERDFHVSHPG 155
Db 121 GSDKLFHW-----GAVKFGKETWSLPNDPDKTKVYKKN-----ALRTPFYK-SG 165
QY 156 KYTLIEVQAEDEAQAENV-----VKVDGPGVIVTKRVQPCSLDPATQKLTNI 203
Db 166 SNSILREIRDTAIEATEFLIYDEAHDKWKIKNNGNFRVKLSRKEIRGPDVSVPEELVQI 225
QY 204 FS-----KEMFKNTMALMDLVKKMPGLKLSKQOIAARGFEALALEEALKGPTDGGQ 255
Db 226 QSYLRWERKQKQNPPEKEEYEAARTVL-----QBEIARG-ASIQDIRARLTKTNDKSO 280
QY 256 SLEELSSHFTVIPHNFHSGQ-----PPPIINSPELLQAKKMDLLVL-ADIEL 301
Db 281 SKEEPLHVTKSDIPDDLAQAQAIYRWEKAGKPNYPPEKQIEELEEAARRELQLEKGITL 340
QY 302 AQALQAVSBOEKTVEEVPPLDRDYQLLKCQLQLDLSG-----APEYKVIQTYLEQT 353
```



